An Electrostatic Spatial Resonance Model for Coaxial Helical Structures with Applications to the Filamentous Bacteriophages

Christopher J. Marzec and Loren A. Day
The Public Health Research Institute, New York, New York 10016 USA

ABSTRACT A model is presented that treats the symmetry matching problem in structures made of two interacting coaxial helices of point charges. The charges are sources of a potential field that mediates a non-specific attractive interaction between the helices. The problem is represented in Fourier space, which affords the most generality. It is found that coaxial helices with optimally mated symmetries can lock into spatial resonance configurations that maximize their interaction. The resonances are represented as vectors in a discrete three-dimensional space. Two algebraic relations are given for the four symmetry parameters of two helices in resonance. One-start inner helices interacting with coaxial one-start or N_R-start outer helices are considered. Applications are made to the filamentous bacteriophages Ff, Pf1, Xf, and Pf3. The interaction given by the linearized Poisson-Boltzmann equation is calculated in this formalism to allow comparison of the electrostatic free energy of interaction of different resonance structures. Experimental nucleotide/subunit ratios are accounted for, and models for the DNA-protein interfaces are presented, with particular emphasis on Pf1.

INTRODUCTION

Specific and nonspecific interactions play different roles in establishing local structure and large-scale symmetry. The inter-dependence of specific interactions and fine scaled local structure is well appreciated. Nonspecific interaction relies on large-scale geometrical features defined over a long distance scale, which therefore are largely independent of small-scale structural features. An example is the binding of DNA by protein, which utilizes features such as the shapes of the major and minor grooves and the spatial distribution of charged groups. The purest instance of the interplay between global symmetry and nonspecific interaction is a macromolecular structure composed of strongly interacting, highly symmetrical substructures. We will argue in this paper that in such a structure, constraints on the symmetries can be determined by optimizing the integrated nonspecific interaction between the sub-structures. Because they are geometrical in origin, these constraints can be studied without a detailed knowledge of the nonspecific interaction that engenders them.

We consider here only the case of cylindrical geometry. An example is the RNA-protein interaction in tobacco mosaic virus, the protein of which forms a cylinder in which a winding of RNA is embedded. Their interaction is non-sequence-specific along its entire length, accommodating any sequence of nucleotides, but always exactly three per protein subunit. This is a particularly simple case of a repeating, nonspecific interaction, because the nucleotide/subunit (n/s) stoichiometry is integral and the RNA and protein share the same helical symmetry. However, other, more complex, nonspecific modes of nucleic acid-protein structural interaction exist, in which the n/s ratio is not necessarily integer, and the symmetries of the nucleic acid and protein substructures are not identical.

The set of filamentous bacteriophages presents a natural theater for studying such nonspecific interactions. These viruses consist of a circular, single-stranded DNA molecule containing several thousands of nucleotides, ensheathed by several thousand copies of a major coat protein subunit, uniformly distributed along the length of the virion; a small number of minor coat proteins are located at the ends (Model and Russel, 1988; Day et al., 1988; Makowski, 1993). In electron micrographs, these virions resemble flexible threads, about 70 nm in diameter and 700–2000 nm in length. Two antiparallel strands of DNA wind around the structure axis, within a cylindrical cavity some 15–20 Å in diameter (Wachtel et al., 1974; Makowski and Caspar, 1978; Day et al., 1988; Reisberg, 1989).

Information concerning the symmetries of the protein coats of the filamentous bacteriophages is from x-ray fiber diffraction studies (Marvin et al., 1974a, b; Nave et al., 1981; Makowski and Caspar, 1981; Peterson et al., 1982). These fall into two groups: Class I, typified by Ff (f1, fd, M13), IKe, and If1, and Class II, typified by Pf1, Xf, and Pf3. It is most significant that within a class the protein coats share the same symmetry. All Class I viruses have fivefold rotational symmetry, with a pentamer axial rise T of about 16 Å and a pentamer rotation angle of about 36°. All Class II viruses have one-start helical symmetry, with axial rise of about 3 Å, protein subunit helix pitch of about 16 Å, and rotation angle of about 67°. Fig. 1 shows lattice diagrams for both types. The origins of these symmetries have been systematically explored in an algebraic model (Marzec and Day, 1988) that builds the virions by close packing α-helices in two layers around a core of DNA, based on the assumption

Received for publication 18 February 1994 and in final form 16 September 1994.

Address reprint requests to Dr. Christopher J. Marzec, Inst. for Biomolecular Stereodynamics, Department of Chemistry, SUNY-Albany, Albany, NY 12222.

0006-3495/94/12/2205/18 $2.00

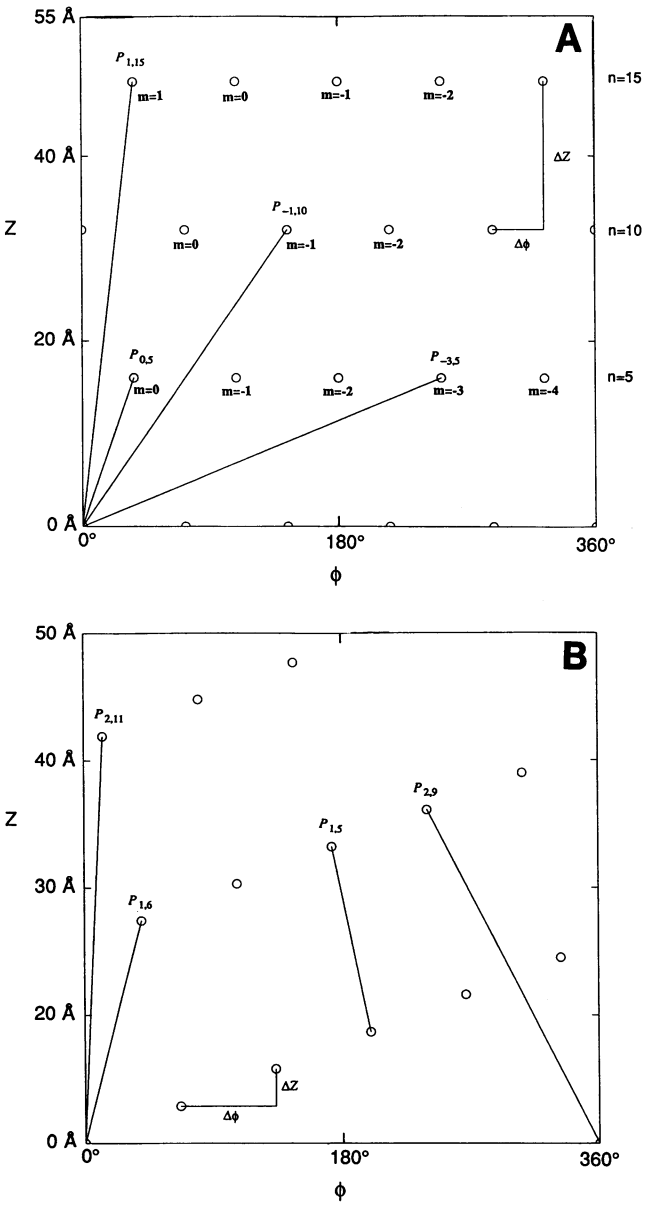

FIGURE 1 (*A*) The Class I surface lattice. The cylindrical coordinates ϕ and z label the position of one arbitrary point of each subunit. The protein coat is organized into levels, each labeled with an index n and spaced along the z axis at intervals of $\Delta z_{pro} = T = 16$ Å; each level has five subunits labeled with indices $-5 < m < 5$. This is a five-start helix with a rotation angle of 36°. The solid lines indicate the directions of the $P_{0,5}$, $P_{-3,5}$, $P_{-1,10}$, and $P_{1,15}$ subhelices; their pitches are $P_{mn} = 10nT/(n - 10m)$. (*B*) The Class II surface lattice. The nth subunit has coordinates $\phi_n = n\,\Delta\phi_{pro}$ and $z_n = n\,\Delta z_{pro}$, where $\Delta\phi_{pro} = 66.66°$ and $\Delta z_{pro} = 2.9$ Å. The $P_{1,6}$, $P_{2,11}$, $P_{1,5}$, and $P_{2,9}$ subhelices are indicated.

that the gently curved α-helices within each layer are fairly straight. This procedure rationalizes the two observed symmetry classes and predicts the possible existence of several others.

Because the DNA contributes only from 6% (in Pf1) to 14% (in Pf3) of the virion mass, it makes a minor contribution to the x-ray diffraction patterns. Most of what is known about the DNA structures derives from spectroscopic and chemical methods, but these do not directly reveal their fundamental symmetries. One exception is a report by Banner et al. (1981) of Ff x-ray fiber diffraction patterns containing two layer lines that they assigned to DNA having a pitch of 26.7 Å; we will consider this result below. Ultraviolet absorbance and circular dichroism spectra indicate that the DNAs of Ff and Xf are similar to classical base-stacked DNAs with bases at the center, but that the DNAs of Pf1 and Pf3 are decidedly nonclassical (Casadevall and Day, 1983; Day et al., 1988; Kostrikis et al., 1994).

The number of nucleotides and the virion length determine the DNA axial rise, Δz_{DNA}. The fundamental stoichiometric quantity n/s is then given by

$$\frac{n}{s} \equiv \frac{2\Delta z_{pro}}{\Delta z_{DNA}} \quad \text{(Class II)} \quad \text{and} \quad \frac{n}{s} \equiv \frac{2T_p}{5\Delta z_{DNA}} \quad \text{(Class I)}. \quad (1)$$

where Δz_{pro} is the axial rise between protein subunits, T_p is the axial rise between pentamers, and the factor of two arises because the DNA has two strands. No single number is more important than *n/s* for describing the nature of the DNA-protein interaction in these structures. Table 1 contains data on the basic symmetries of the DNA cores and the protein coats of the filamentous bacteriophages. The experimental *n/s* value for Pf1 is definitely the integer one; those of Ff and its mutants and those of Pf3 are definitely non-integer; and that of Xf is probably noninteger.

For these types of structure, there is no generally applicable direct technique for determining either the rotation angle of the DNA helix, $\Delta\phi_{DNA}$, or equivalently, the pitch of the DNA helix, P_{DNA}, so even the basic symmetries of the DNA helices are not easily accessible. The ^{31}P NMR data indicate that the Pf1 DNA backbone has only one phosphate group orientation, whereas that of Ff shows many (Cross et al., 1983). These data suggest that for Pf1, at least, the DNA is packed helically, because a DNA molecule structured otherwise would have a very large number of chemical environments. Also, the x-ray diffraction patterns of Ff reported by Banner et al. (1981) imply that the DNA is in a regular helix that, the NMR data indicate, must have many different environments. We assume for this paper that the DNA molecule in a filamentous virus is indeed packed in a regular helix.

We will further assume that the DNA and protein coats are strongly related. No filamentous bacteriophage has ever been observed as a "ghost," suggesting that a protein-DNA interaction is at least a necessary part of the assembly process. Hunter et al. (1987) performed site-directed mutagenesis experiments on Ff, altering one of the four lysine residues (Lys^{48}, Table 1) in its DNA interaction domain to any of three uncharged residues, reducing the possible charge at the C-terminus from +4 to +3. This had the effect of increasing the virion length by 33–38%, presumably by causing the DNA axial rise to increase to maintain charge balance with the protein. Nature has presented a similar, but more extensive experiment involving Pf1 and Xf: despite the evolutionary linkage shown by the high homology of their protein coats, their DNA axial rises differ by about a factor of two,

TABLE 1 Values for n/s (nucleotides per major coat protein subunit)

Virus		Amino acid sequence of coat protein subunit	n/s Values
CLASS I			
Ff (fd)	(a)	AEGDDPAKAAFDSLQASATEYIGYAWAMVVVIVGATIGIKLFKKFTSKAS	2.38 ± 0.05 (a1) 2.35 ± 0.06 (a2) 2.33 ± 0.14 (a3) 2.38 ± 0.08 (a4) 2.41 ± 0.14 (a5) 2.35 ± 0.05 (a6) 2.29 ± 0.04 (a7)
fd-K48Q	(b)	AEGDDPAKAAFDSLQASATEYIGYAWAMVVVIVGATIGIKLFKKFTSQAS	1.80 (b1)
fd-K48T		AEGDDPAKAAFDSLQASATEYIGYAWAMVVVIVGATIGIKLFKKFTSTAS	1.73 (b1)
fd-K48A		AEGDDPAKAAFDSLQASATEYIGYAWAMVVVIVGATIGIKLFKKFTSAAS	1.76 (b1)
CLASS II			
Xf	(c)	SGGGGVDVGDVVSAIQGAAGPIAAIGGAVLTVMVGIKVYKWVRRAM	2.07 ± 0.15 (c1) 2.12 ± 0.15 (c2) 2.16 ± 0.09 (c3) 2.15 ± 0.10 (c4)
Pf1	(d)	GVIDTSAVQSAITDGQGDMKAIGGYIVGALVILAVAGLIYSMLRKA	0.97 ± 0.05 (d1) 1.01 ± 0.02 (d2) 1.05 ± 0.06 (d3) 1.00 ± 0.07 (d4) 1.01 ± 0.04 (d5)
Pf3	(e)	MQSVITDVTGQLTAVQADITTIGGAIIVLAAVVLGIRWIKAQFF	2.38 ± 0.14 (e1) 2.57 ± 0.08 (e2) 2.40 ± 0.09 (e3) 2.40 ± 0.04 (e4)

For sequences: (a) Beck and Zink (1981); (b) Hunter et al. (1987); (c) Day et al. (1988); (d) Hill et al. (1991); (e) Luiten et al. (1985).

For stoichiometries: (a1)* N/P ratio of 27.6 (Hoffmann-Berling et al., 1963); (a2)* g P/g virus and (a3)* spectral regression (Berkowitz and Day, 1976); (a4)* total mass, Svedberg eqn (Newman et al., 1977); (a5) mass-per-length, STEM (Reisberg, 1989), (a6) and (a7) $2\Delta z_{DNA}/\Delta z_{pro}$ from c-spacings and lengths for dry and wet virus, data from (Frank and Day, 1970; Dunker et al., 1974; Newman et al., 1977).

(b1) Estimated from length data from Hunter et al. (1987) and an n/s = 2.40 for fd.

(c1)‡ g P/g virus (Wiseman and Day, 1977), corrected for sequence (c); (c2) mass-per-length, STEM (Reisberg, 1989) (c); and (c3) $2\Delta z_{DNA}/\Delta z_{pro}$ from c-spacing and EM length (wet virus) (Marvin et al., 1974; Chen et al., 1980).

(d1)‡ N/P mol ratio, and (d2)‡ g P/g virus (Wiseman and Day, 1977); (d3) mass-per-length, STEM (Reisberg, 1989); (d4) spectral regression (Kostrikis et al., 1994); (d5) and (d6) $2\Delta z_{DNA}/\Delta z_{pro}$ from c-spacing and EM length (dry virus) (Marvin et al., 1974; Wiseman et al., 1976).

(e1)‡ Total mass, Svedberg Eq. (Newman et al., 1982); (e2) mass-per-length, STEM (Reisberg, 1989); (e3) and (e4) $2\Delta z_{DNA}/\Delta z_{pro}$ from c-spacings and EM lengths (wet and dry virus) (Peterson et al., 1982; Newman et al., 1982; Reisberg, 1989).

*Values for fd so marked have been corrected on the basis that 33,000 daltons from g7p and g9p can be grouped with g8p in the n/s calculation, but 300,000 daltons of protein mass from g6p and g3p is excluded.

‡Values for Class II viruses so marked have not been corrected for minor components as in *; such corrections would increase these apparent n/s values by about 2% for Xf and Pf3, and about 1% for Pf1.

and the Xf protein subunit contains four positive charges at its C terminus, compared with two for Pf1 (Table 1). Sedimentation coefficients of Pf3-Ag^+ complexes are larger than the values for the native virion, and the magnitude of the increase is greater than can be accounted for by the mass of the Ag^+ ions alone. Because the Ag^+ ion binds to the DNA, this suggests that some structure modification of the DNA propagates outward through the protein coat (Casadevall and Day, 1983). These observations on the four viruses treated herein strongly suggest that the DNA and protein are structurally related and that the controlling part of their interaction is electrostatic.

Averaged over its length, a complex of DNA and protein with nonrelated symmetries would interact only weakly through their average properties; a stronger DNA-protein interaction requires some symmetry matching arrangement. TMV solves this problem easily, by winding its RNA in a helix of 4 nm radius, allowing it to follow the protein coat exactly. In general, however, the problems of symmetry matching are not trivial. How can the rotationally symmetrical protein coats of the Class I viruses mate in a straightforward way with the helical DNA symmetry? The Class II viruses have helically symmetrical protein coats, but what happens if the pitch of the protein helix, P_{pro}, does not equal the pitch of the DNA helix, P_{DNA}? And even if these pitches are equal, and the helices can follow each other, how does one understand a nonintegral *n/s* value?

In an earlier approach to these questions, we considered the DNA-protein interaction in essentially mechanical terms, positing a ridges-into-grooves arrangement in which the outer parts of the DNA molecule fitted into grooves on the inner surface of the protein sheath (Marzec and Day, 1983). These grooves were created by the undulating surface of the α-helices forming the C terminus of the protein subunits, one α-helix per subunit, so arranged that they surround the DNA in *N* continuous tubes winding up the structure axis; where the α-helix tube of one subunit leaves off, another begins. Such a mechanical interconnection between symmetries, which we dubbed the "pitch connection," probably does contribute to the DNA-protein interaction. However, only a part

of the DNA-protein interaction is mechanical, caused by shapes and matched surfaces. In this paper, we consider electrostatic interactions between the negatively charged DNA and the positively charged C-terminal region of the protein coat. We show that the posited electrostatic interaction produces a more general version of the mathematical relationships between DNA and protein helix parameters. They emerge naturally from a mathematical description of maximally interacting coaxial helices, but require no assumptions about the structure of the protein C-terminus, except that it is positively charged. However, we will show that input of these positions, as well as the inclusion of two strands of DNA, only modifies the strengths of the maximal interactions, introducing no change in the algebraic relationships among the DNA and protein helix parameters that correspond to maximal interaction between a single DNA helix and a coaxial one-start (for Class II) or five-start (for Class I) helix.

THEORY

Our working hypothesis is that the DNA molecule adopts helix parameters that maximize the electrostatic interaction with the protein sheath. The measure of this interaction is the Helmholtz free energy, $A = E - TS$. It might seem that nothing useful can be said about the electrostatic free energy without detailed knowledge of the electrostatic nature of the DNA-protein interface, but this is not the case. The electrostatic field created by symmetrically arranged charged residues is smoother and longer-ranged than the local fields associated with chemical interactions, so on the nanometer scale, it tends to blur local details, while retaining long-scale structural information. We show in this section that relationships between the symmetries of interacting coaxial helixes can be determined by optimizing the geometry of the interaction. The details surely affect the magnitude of the interaction, but they are entirely irrelevant to the sought after relationships. In this view, the DNA-protein interaction is characterized by a relatively long spatial scale that obliterates the effects of local variations in components, while knitting them into a highly symmetrical structure.

The helically symmetrical potential field

We begin by considering the functional form of the electrostatic potential field created by helix 1, an infinite helix of point charges with axial rise Δz_1 and rotation angle $\Delta\phi_1$. This form is derived in the Appendix:

$$V_1(r, \phi, z) = \sum_{m=-\infty}^{m=+\infty} \sum_{n=-\infty}^{n=+\infty} A_{mn}(r)\exp[in(\phi - \alpha)]\exp[-ik_{mn}^{(1)}(z - \beta)], \quad (2)$$

where

$$k_{mn}^{(1)} \equiv \frac{n\Delta\phi_1 - 2\pi m}{\Delta z_1}. \quad (3)$$

The form of $k_{mn}^{(1)}$ is already familiar from the selection rules for the layer lines appearing in diffraction from fibers of helical molecules: $l/c = n/P - m/\Delta z = (1/2\pi)k_{mn}^{(1)}$. (The diffracting discontinuous helix has pitch P and axial rise Δz; the sign of m has been changed for consistency with $k_{mn}^{(1)}$ but this is without physical consequence because all integers m and n are used to calculate the possible l/c values.) Equation 2 is simply a Fourier decomposition of the ϕ and z dependence of V_1. It is clear from inspection that V_1 fulfills the condition of helical symmetry: $V_1(r, \phi + \Delta\phi_1, z + \Delta z_1) = V_1(r, \phi, z)$. The constants α and β are arbitrary offsets used below, and r is the distance from the z axis. The coefficients $A_{mn}(r)$ determine the nature of the field V_1, and they are evaluated in the Appendix from linearized Poisson-Boltzmann (LPB) electrostatics. The Appendix also gives the form for a potential field with N_r-fold rotational symmetry, a slightly modified version of Eqs. 2 and 3 in both of which index n is replaced by sN_r, and the sum in modified Eq. 2 is over all s values. The form of Eq. 2 applies to all helically symmetrical potential fields so, quite generally, it describes the field created by any helically symmetrical distribution of sources; in the electrostatic case, these are point charges, dipoles, or higher multipoles. If the source helix has radius r_s, then we can write $A_{mn} = A_{mn}(r, r_s)$.

Spatial resonance

Now consider a second helix of point charges q_j with cylindrical coordinates: $[r_2, j\Delta\phi_2 + \gamma, j\Delta z_2 + \delta]$, where $j = -J$ to $+J$, and γ and δ are arbitrary offsets. Define the electrostatic free energy per unit length:

$$A/L = \left(\frac{1}{L}\right) \sum_{j=-J}^{J} V_1(\vec{r}_j)q_j, \quad (4)$$

where helix 2 has $N = 2J + 1$ charges within a distance $L = (2J + 1)\Delta z_2$ along the common structure axis. This is the entire electrostatic free energy of interaction between the two helices. The usual factor of (½) does not appear because the free energy of the charges of helix 2 in the potential field created by the charges of helix 1 is equal to the free energy of the charges of helix 1 in the potential field created by the charges of helix 2; instead of doing both sums and applying the factor (½), we do one sum explicitly and omit the (½). We will set $q_j = Q$ for all j. Thus,

$$\frac{A}{L} = \left(\frac{Q}{L}\right) \sum_{j=-J}^{J} \sum_{m=-\infty}^{m=+\infty} \sum_{n=-\infty}^{n=+\infty} A_{mn}(r_2, r_1) \times \exp\{i[n(j\Delta\phi_2 + \gamma - \alpha) - k_{mn}^{(1)}(j\Delta z_2 + \delta - \beta)]\} \quad (5)$$

$$= \left(\frac{Q}{L}\right) \sum_{m=-\infty}^{m=+\infty} \sum_{n=-\infty}^{n=+\infty} B_{mn} \sum_{j=-J}^{J} \exp\{ij[n\Delta\phi_2 - k_{mn}^{(1)}\Delta z_2]\},$$

where we have absorbed some of the exponential factors and A_{mn} into B_{mn}:

$$B_{mn} \equiv A_{mn}(r_2, r_1)\exp\{i[n(\gamma - \alpha) - k_{mn}^{(1)}(\delta - \beta)]\}. \quad (6)$$

The j summation may be done through the relation

$$\sum_{j=-J}^{J} e^{ijx} = \frac{\sin[(J + 1/2)x]}{\sin(x/2)}. \quad (7)$$

If $x = 2\pi n$, the right-hand side of this equation attains its maximum possible value, $2J + 1$. Thus, from Eqs. 5 and 7,

$$\frac{A}{L} = \left(\frac{Q}{L}\right) \sum_{m=-\infty}^{m=+\infty} \sum_{n=-\infty}^{n=+\infty} B_{mn} \frac{\sin[(J + 1/2)(n\,\Delta\phi_2 - k_{mn}^{(1)}\Delta z_2)]}{\sin[(n\Delta\phi_2 - k_{mn}^{(1)}\Delta z_2)/2]}. \quad (8)$$

This equation shows the resonance character of the interaction free energy. The ratio of sines is maximum for the M, and N term when $N\Delta\phi_2 - k_{MN}^{(1)}\Delta z_2 = 2\pi K$ for some three integers M, N, and K; or, substituting for k_{MN},

$$\frac{N\Delta\phi_1 - 2\pi M}{\Delta z_1} = \frac{N\Delta\phi_2 - 2\pi K}{\Delta z_2} \quad \text{or} \quad k_{M,N}^{(1)} = k_{K,N}^{(2)} \quad (9)$$

for the three integers (K, M, N). Let the triple (K, M, N) represent a vector $\vec{R}$ in a three-dimensional lattice space, which consists of the lattice of points with integer values of K, M, and N. Then we will refer to a resonance as $\vec{R} = (K, M, N)$.

The largest contribution to A/L is from the term $(Q/L) \cdot (2J + 1)B_{MN} = (Q/\Delta z_2)B_{MN}$. Thus, if $\Delta\phi_1$, Δz_1, and $\Delta\phi_2$ are fixed, the resonance condition may be encountered for values of Δz_2 that obey Eq. 9. The resonance becomes sharper as J (or L) is increased, and as J (or L) approaches infinity, it becomes infinitely sharp, so in that case A/L would vanish unless the resonance condition holds exactly for some set (K, M, N). If helix 1 and helix 2 represent the protein and DNA helices of a filamentous phage, then both helices are finite, with J values of several thousand. If both helices are taken to be finite, the resonance condition of Eq. 9 is unaffected, but the resonance is somewhat less sharp.

We can better understand the spatial nature of the resonance by introducing the "subhelix" pitches $P_{mn}^{(1)}$ defined for lattice 1 in the Appendix:

$$P_{mn}^{(1)} \equiv \frac{2\pi n}{k_{mn}^{(1)}} = \frac{2\pi n\Delta z_1}{n\Delta\phi_1 - 2\pi m} \quad (10)$$

describes one start (Class II) helices; for N_R-start helices ($N_R = 5$ for Class I virions), n is a multiple of N_R and Δz equals the z spacing between N_R-mers. Fig. 1 shows examples of subhelices of Class I and Class II protein lattices. The $m/n = 2/11$ subhelix of the Class II lattice denotes the backbone of the C-terminal portion of the Pf1 subunit (Makowski and Caspar, 1981); the subhelices and the M/N notation appear prominently in a previous algebraic modeling study of the filamentous phages (Marzec and Day, 1988). Any pair of integers m and n and a set of lattice parameters have a corresponding subhelix, which can be located on a lattice diagram and denotes a direction on the lattice diagram. This notation is more compact than writing, for example, "the direction of the 6-start helix," meaning the $P_{1,6}$ helix; the verbal description is inadequate for higher order subhelices, such as the $P_{3,17}$ helix. Klug et al. (1958) introduced these helices in the context of fiber diffraction theory, referring to helical projection down the helix with pitch nc/l, which in our notation is P_{mn}.

Writing $P_{kn}^{(2)} \equiv 2\pi n\Delta z_2/(n\Delta\phi_2 - 2\pi k) = 2\pi n/k_{mn}^{(2)})$ for lattice 2, the resonance condition of Eq. 9 becomes

$$P_{MN}^{(1)} = P_{KN}^{(2)}. \quad (11)$$

(If one helix has N_R-fold rotational symmetry, then in this equation N must be a multiple of N_R.) Thus, a "single spatial resonance" occurs when the protein subhelix of pitch $P_{MN}^{(1)}$ associated with helix 1 is identical to the DNA subhelix of pitch $P_{KN}^{(2)}$ associated with helix 2; i.e., the DNA and protein lattices share a subhelix. The vector $\vec{R}$ labels the shared subhelices with pitches $P_{M,N}^{(pro)} = P_{K,N}^{(dna)}$. If $K = 0$, then the subhelix $P_{O,N}^{(2)}$ of lattice 2 has pitch $2\pi\Delta z_2/\Delta\phi_2$, which is the pitch of the basic DNA helix; similarly for $M = 0$ and the basic protein helix of lattice 1. A resonance that utilizes one of the basic helices is likely to be strong. Equation 11 or 9 determines one relationship between the four helix parameters.

Double resonance

A "double resonance" condition obtains if two triples of integers, say $\vec{R}_1 \equiv (K_1, M_1, N_1)$ and $\vec{R}_2 \equiv (K_2, M_2, N_2)$, satisfy the resonance condition. We then have $P_{M_1,N_1}^{(1)} = P_{K1,N1}^{(2)}$ and $P_{M_2,N_2}^{(1)} = P_{K_2,N_2}^{(2)}$. We will denote this solution as $\vec{R}_1 \,\mathrm{X}\, \vec{R}_2$. If $\vec{R}_1$ and $\vec{R}_2$ are resonances, then any linear combination with $\vec{R}_{pq} \equiv p\vec{R}_1 + q\vec{R}_2$, for integers p and q, will work as well; ie., $K = pK_1 + qK_2$, etc. The tips of the vectors $\vec{R}_{p,q}$, evaluated for all integers p and q, determine a two-dimensional net of points, the plane that passes through $\vec{R}_1$, $\vec{R}_2$, and the origin, $K = M = N = 0$. Thus, the two resonance vectors $\vec{R}_1$ and $\vec{R}_2$ generate an infinite plane of resonances in KMN space, a double resonance "family." Although every point in the KMN lattice space corresponds to a single possible resonant interaction between two coaxial helices, only helices in double resonance share an entire plane of resonant interaction points.

A plane through the origin is determined uniquely by its normal vector, so we define the "double resonance vector"

$$\vec{D} \equiv \vec{R}_1 \,\mathrm{X}\, \vec{R}_2, \quad (12)$$

where the X denotes the usual vector cross product, and $\vec{D}$ has components: $D_1 = M_1N_2 - M_2N_1$, $D_2 = N_1K_2 - N_2K_1$ and $D_3 = K_1M_2 - K_2M_1$. The ratios of the components of a vector $\vec{D}$ are rational numbers, so each $\vec{D}$ labels a lattice point in the KMN space. We will denote a double resonance family by its normal vector $\vec{D}$; the vector $c\vec{D}$, for any real, non-zero constant c, indicates the same double resonance. The components of $\vec{D}$ will be separated by slashes, i.e., $(D_1/D_2/D_3)$, to avoid confusion with the components of a resonance $\vec{R}$, which will be separated as above with commas. Each double resonance $\vec{D}$ has associated with it a plane lattice formed from the tips of all resonance vectors $\vec{R}_{p,q}$, but we will see below that a few of these are much more significant than all the others. If linear combinations of $\vec{R}_1$ and $\vec{R}_2$ are used to

calculate $\vec{D}$, its direction remains unchanged and only its magnitude is affected. Given a double resonance $\vec{D}$, basis vectors for the double resonance plane can be written

$$\vec{r}_1 = [0, D_3, -D_2]$$
$$\vec{r}_2 = [-D_3, 0, D_1] \quad (13)$$
$$\vec{r}_3 = [D_2, -D_1, 0].$$

Any integer sums of these generate possible double resonance vectors $\vec{R}$. Because $0 = D_1\vec{r}_1 + D_2\vec{r}_2 + D_3\vec{r}_3$, these vectors are not linearly independent.

The double resonance condition gives two Eqs. 9 in four variables Δz_{pro}, $\Delta\phi_{pro}$, Δz_{DNA}, and $\Delta\phi_{DNA}$, with solutions

$$0 = D_1\Delta\phi_{DNA} + D_2\Delta\phi_{pro} + 2\pi D_3 \quad (14)$$

and

$$0 = D_1\Delta z_{DNA} + D_2\Delta z_{pro}. \quad (15)$$

Equations 14 and 15 are our main result. These equations show that the direction of $\vec{D}$ is meaningful, but not its magnitude or its sign. Helix parameters related as above by a double resonance condition correspond to an infinite family of shared subhelices; examples of double resonances will be seen in the applications.

Recasting Eqs. 14 and 15 reveals other relationships. Equation 15 combined with Eq. 1 gives

$$\frac{n}{s} = \frac{-2D_1}{D_2}(1 - \text{start})$$

and (16)

$$\frac{n}{s} = \frac{-2D_1}{N_R D_2}(N_R - \text{start}),$$

where $N_R = 5$ for the Class I viruses. Dividing Eq. 14 by Δz_{DNA} and using Eq. 15 yields

$$0 = \frac{1}{P_{DNA}} - \frac{1}{P_{pro}} + \frac{D_3}{L_{repeat}}, \quad (17)$$

where $L_{repeat} \equiv D_1\Delta z_{DNA} = -D_2\Delta z_{pro}$ is the axial repeat distance for the DNA-protein interaction. This equation has the same form as the "restricted pitch connection" equation (Marzec and Day, 1983), but with a third term that reflects the more general nature of the present considerations.

Every direction in *KMN* space that intersects one of its lattice points corresponds to some possible double resonance $\vec{D}$. A structure in a given double resonance configuration is described by one and only one $\vec{D}$; a different $\vec{D}$ describes a different structure. To compare two double resonance structures, $\vec{D}_1$ and $\vec{D}_2$, numerical values for the B_{MN} are needed. Comparison is important because Eq. 16 shows that all rational numbers correspond to double resonance n/s values, which is clearly not a useful result unless some are much more powerful than others. However, the value of A/L can remain finite only if the sum over the A_{MN} values of a family converges, and this suggests that only for small integers M and N will the A_{MN}, and hence the B_{MN}, be large. The finitude of the actual helices is also helpful, because it disallows the ultra-long scale resonances corresponding to very large integers.

The four-zone model

To render these observations quantitative, a physical model for the electrostatic interaction is needed. A rough guide to the size of the B_{MN} values is sufficient. This is fortunate, because the electrostatics problem represented by a DNA-protein interface cannot be solved exactly. The distance between the DNA and protein surfaces is only several angstroms; solvent molecules may be present; the dielectric properties of the DNA and protein are different, and dielectric behavior is poorly defined at such small distances. Also, a laboriously obtained more accurate solution would be of little use, because we wish to examine large ranges of helix parameters. Thus, we will content ourselves with a solution to the LPB equation, which will give a reasonable, qualitatively useful guide to the sizes of the double resonances.

The Appendix calculates, in the LPB approximation, the electrostatic potential field caused by a helix of point charges. This calculation amounts to determining the A_{mn} values, which emerge as an appropriately weighted product of modified Bessel functions. This potential field has been presented in a different form by Soumpasis (1978) to extend the range of the Debye-Huckel condensation theory of Manning. His expression involves only the K_0 Bessel function, which is computationally convenient; however, its form does not allow for easy calculation of the resonant interaction with a second helix or for continuous variation of helix parameters.

For the most plausible simple model consistent with the LPB equation, we have divided the space into four coaxial regions, each of which can be given its own dielectric constant ϵ and Debye damping factor κ: region I is the innermost, from $r = 0$ to $r = r_{DNA}$, where r_{DNA} represents the outer radius of the (presumed roughly cylindrical) DNA molecule; region II is the innerzone, between r_{DNA} and r_{pro}, the innermost edge of the protein coat; region III represents the (presumed cylindrical) protein coat, from $r = r_{pro}$ to $r = r_V$, the virion radius, about 30–35 Å. The double sum converges slowly, but we find for $|r_{pro} - r_{DNA}| > \sim 0.5$ Å that including terms beyond $n = 50$ causes fractional changes of less than 10^{-3}. A simpler model with constant ϵ and κ values would also allow computation of the A_{mn} values, but the more elaborate calculation allows us to assess to what extent the relative strengths of the A_{mn} depend on the details. If it were found that changes in the ϵ or κ values alter the relative strengths of the A_{mn}, then it would be far more difficult to locate the most significant double resonances. What we do find upon varying the model parameters is that the relative strengths of the A_{mn} do not change significantly.

So far we have considered only the DNA-protein interaction, neglecting the self-energies of the DNA and protein helices. The DNA-self energy arises from both inter-strand and intra-strand interactions. Because of the large distances

between charges in usual DNA geometry, as is probable in Xf, IKe, If1, and Ff (Day et al., 1988), self-energy contributions are minor compared with the DNA-protein interaction. However, in the special case of Pf1, discussed in detail below, the DNA self-energy contributes significantly in determining the virion structure. To demonstrate this numerically, the total self-energies of the DNA strands and of the protein charges have been calculated and added to the DNA-protein interaction energy. This is easy to do in the simplified case of κ and ϵ values assumed to be constant in all four zones, because then the contribution from each charge is simply $qe^{-\kappa r/\sqrt{\epsilon}}/(r\epsilon)$. We find that generally the self-energy is negligible compared with the interaction energy. (We have calculated the DNA-protein interaction energy in this way for a uniform medium, for comparison to the four-zone model; the two calculations agree to within 1%, for DNA and protein charge separations greater than ½ Å, with errors caused by series truncation.) Because our purpose is to understand the DNA-protein interaction, we will generally omit consideration of the DNA and protein self-energies, except in the special case of Pf1.

Distributed charges, the product lattice, and resonant phasing

Although the relative sizes of the A_{mn} values are insensitive to details, the situation is different with the B_{mn}, because the phase factors, depending on the relative offsets of the DNA and protein helices, can dramatically change the value of the B_{mn}. Here we introduce the facts that there are two strands of DNA, and that the protein subunits typically contain from two to four basic residues (see Table 1). Let the coordinates of charges of the 0 th subunit be r_t, α_t, β_t, where $t = 1$ to T; treat the two DNA strands as a single helix of pairs of nucleotides, containing two charges per nucleotide pair, and put the DNA charges in the 0th residue at coordinates: r_s, γ_s, δ_s, where $s = 1$ and 2. These charges can be introduced by summing the contributions from all the basic residues in finding V_1, and by summing over both strands of DNA when calculating the free energy. Thus,

$$B_{mn} \equiv \sum_{t=1}^{T} e^{-i[n\alpha_t - k^{(1)}_{mn}\beta_t]} \sum_{s=1}^{2} e^{i[n\gamma_s - k^{(1)}_{mn}\delta_s]} A_{mn}(r_s, r_t). \quad (18)$$

These sums are form factors for the charge distributions, written for helical geometry. We will suppose that the structure is in a double resonance configuration, so we need to evaluate the effect of the offsets on B_{mn} only when (m, n) belong to a double resonance family. The maximum values of the first and second sums are T and 2, respectively, which occurs if every square bracket of Eq. 18 equals a multiple of 2π. These equations are the same for the protein offsets (α, β) and for the DNA offsets (γ, δ), so denoting the angular offsets by ω and the z offsets by λ, we have

$$N_1\omega - k^{(1)}_{M_1, N_1}\lambda = 2\pi L_1$$

and (19)

$$N_2\omega - k^{(1)}_{M_2, N_2}\lambda = 2\pi L_2.$$

The solutions ω, λ form a lattice, which we will call the "product lattice," indexed by the integers L_1 and L_2. (The product lattice exists in the two-dimensional ϕ-z space, and it should not be confused with the *KMN* lattice space or the planar lattice of points in the *KMN* space that correspond to a double resonance.) If some numbers ω and λ solve Eqs. 19 for two members of a double resonance family $\vec{R}_1 = (K_1, M_1, N_1)$ and $\vec{R}_2 = (K_2, M_2, N_2)$ and two integers L_1 and L_2, then they will solve the equivalent equation for any resonance in that family, $\vec{R}_{pq = p\vec{R}_1 + q\vec{R}_2}$ and $L = pL_1 + qL_2$. Thus, a given product lattice corresponds to a particular double resonance $\vec{D}$. Equations 19 are solved for $(\omega, \lambda) = (\Delta\phi_{pro}, \Delta z_{pro})$ or, using Eq. 9, for $(\omega, \lambda) = (\Delta\phi_{DNA}, \Delta z_{DNA})$. So the product lattice contains the basic lattices of the DNA and protein; the product lattice for a Class I protein coat shares its fivefold rotational symmetry.

The product lattice can be understood geometrically as well. One can suppose, for the moment, that the subunits contain a single charge and set the protein charge of the 0th subunit directly "above" (same ϕ and z, greater r) the DNA phosphate charge of the 0th nucleotide of one DNA strand. The DNA helical symmetry then fixes the positions of the charges of all other nucleotides in the strand, generating a single DNA helix on the ϕ-z plane. This procedure can be repeated for a series of imaginary DNA strands, each positioned with its 0th nucleotide charge under a different protein helix charge, until all of the protein charges have a DNA charge under them. The resulting set of DNA helices then forms the product lattice.

If the protein subunit has T charges, another $T - 1$ charges can be added to the 0th subunit, using $T - 1$ sets of offsets ω and λ, which are solutions of Eqs. 19, so that all T charges lie on the product lattice. Then the protein helical symmetry places all the charges of all T helices in optimum positions on the product lattice. The same can be done for the charges of the second DNA strand, giving two DNA helices on the product lattice. We will call a helix of charges positioned on the product lattice "resonantly phased," so this construction describes $T + 2$ resonantly phased helices. From the standpoint of the DNA-protein interaction, it makes no difference which solutions of Eqs. 19 are used (which helical sets of product lattice points are occupied), because averaged over the virion, all resonantly phased helices are equivalent. Figs. 3, 7, and 9, discussed below, show examples of product lattices.

Optimized double resonance

The double resonance conditions of Eqs. 14 and 15 address the problems mentioned above by determining a relation between two coaxial helices. But a structure in a given double resonance, $\vec{D}$, still has two free parameters. Because n/s is fixed by Eq. 16 one of these two parameters can be an axial rise, and because $\Delta\phi_{DNA}$ is given in terms of $\Delta\phi_{pro}$ by Eq. 14, the other can be a rotation angle. Taking the free parameters to be Δz_{pro} and $\Delta\phi_{pro}$, we now ask whether the interaction function can be optimized in these as well. Because the

double resonance conditions were defined by requiring that the exponential factor of Eq. 5 be unity, the net strength of the interaction is the sum of the A_{MN} values in the double resonance family. We will seek to optimize this sum.

Because n/s is fixed for a given double resonance, varying Δz_{pro} automatically causes Δz_{DNA} to adjust in concert, so the net effect of changing Δz_{pro} is to redefine the length scale of the problem, without altering its angular geometry. A change in the length scale amounts only to a change in the radii of the two helices, so a monotonic increase in Δz_{pro} produces a monotonic change in the net interaction, without giving local minima in the free energy. Thus, it is not possible to optimize the interaction function in Δz_{pro}.

The situation regarding $\Delta\phi_{pro}$ is more interesting. It is shown in the Appendix that in the LPB model, the M and N dependence of the A_{MN} values arises through $(k_{MN}^2 + \kappa^2)^{1/2}$, which appears as the argument of modified Bessel functions. This shows that: A_{MN} is maximum for a $\Delta\phi_{pro}$ value that causes k_{MN} to vanish; the maximum value of A_{MN} decreases with N; and a small value of Δz_{pro} sharpens the dependence on $\Delta\phi_{pro}$, whereas a large value washes out the maximum. The Appendix also proves that this behavior is fairly general, arising from any spherically symmetrical interaction potential that falls off with distance from its source, in a uniform medium. The result is that the value $\Delta\phi_{pro} = 2\pi M/N$ gives a maximum in A_{MN}, even for a more general interaction than that of the LPB equation. Thus, of the infinite lattice points of a given double resonance family, those two or three with the smallest N values dominate the sum over m and n for the interaction energy, causing a local minimum in A/L when k_{MN} vanishes at $\Delta\phi_{pro} = 2\pi M/N$. Because the double resonance has $k_{MN}^{(pro)} = k_{KN}^{(dna)}$, such an optimized interaction fixes $\Delta\phi_{DNA} = 2\pi K/N$ as well. We will give examples below.

APPLICATIONS

Our purpose in this section is not to produce detailed structure models of any of the filamentous bacteriophages, but to demonstrate how the double resonance hypothesis might illuminate the study of their DNA-protein interfaces. As a working hypothesis, we have assumed that the pair of coaxial helices corresponding to each virion is near an electrostatic equilibrium configuration, so that there is no large imbalance of electrical forces which must be offset by other forces. Combined with the double resonance model, this gives the "double resonance hypothesis": structures containing pairs of coaxial helices have symmetry parameters corresponding to an optimized double resonance. Within this framework, each virion entails a separate analysis, depending on available data and its own idiosyncracies.

Ff

Because it has been the subject of x-ray diffraction and mutagenesis studies, more is known about Ff than about other Class I viruses. Its Class I protein lattice has a rotation angle of ~36°; the Class I axial rise of ~16 Å, combined with the Ff n/s value of ~2.4, give Δz_{DNA} ~2.66 Å for Ff. If the two layer lines reported for Ff DNA (Banner et al., 1981) can be taken at face value, they indicate a DNA pitch of 26.7 Å. Combined with Δz_{DNA}, this yields a DNA rotation angle $\Delta\phi_{DNA} = 36°$ (see Marzec and Day, 1983). The mutant Ff form has n/s ~1.8, and its overall morphology (e.g., virion mass per length) is unchanged, but there is no direct data concerning its DNA rotation angle.

The LPB model allows us to calculate the interaction energy for a finite stretch of nucleotides in the potential field created by the protein, as a function of arbitrary helix parameters. We can visualize double resonances via contours on the n/s-$\Delta\phi_{DNA}$ plane, as shown in Fig. 2; it was prepared for a Class I protein lattice, with fivefold rotational symmetry, based on a single charge per subunit and a single DNA helix. If the other DNA strand and other protein charges are included on resonantly phased helices, then the resulting contour plots are identical to a plot made with only one-charge basic helices, except for the scaling factor of $2T$. Each minimum in the interaction energy contour plot corresponds to some double resonance $\vec{D}$. The figure has fivefold symmetry in $\Delta\phi_{DNA}$, with mirror symmetry around the lines $\Delta\phi_{DNA} = (36\,j)°$. Only contours of the deepest minima are displayed, and these double resonances can be compared with the parameters of the native and mutant Ff virions.

The simplest way to determine the double resonance corresponding to any given minimum, some $(n/s, \Delta\phi_{DNA})$ point, is to scan through all possible $\vec{D}$ vectors with small components, sifting out those which, when combined with the Class I protein lattice parameters through Eqs. 14 and 15, yield Δz_{DNA} and $\Delta\phi_{DNA}$ sufficiently near the given target values. The n/s value of 2.4 and the assumption that the DNA

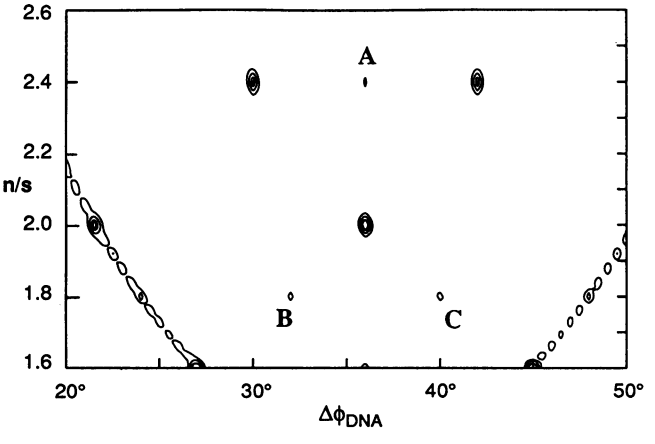

FIGURE 2 The interaction energy contour plot for different DNA helices coaxial with a Class I protein coat, with DNA rotation angle $\Delta\phi_{DNA}$ on the abscissa and n/s on the ordinate. The interaction energies were calculated based on the four-zone model, with $\epsilon = 3$, $\kappa = 0.5$ Å^{-1}, the DNA charges at 10 Å radius, and the protein charges at 11 Å radius. The choices of the actual values for ϵ, κ, and the radii do not change the positions of any of the spots and have only small effects on their relative strengths. The interaction was calculated for a 50 residue DNA molecule. The native Ff double resonance is labeled A, with $n/s = 2.4$ and $\Delta\phi_{DNA} = 36°$; possible solutions for the mutant Ff, with $n/s = 1.8$, are labeled B and C.

pitch is 26.7 Å give, for the native Ff structure, $\Delta\phi_{DNA} = 36°$, which corresponds to a moderately strong double resonance: $\vec{D}_{Ff(native)} = (-60/10/5)$ (point *A* on Fig. 2). The mutant phage has two possible double resonances in the vicinity of $n/s = 1.8$ and a DNA rotation angle near 36°: the minimum at $\Delta\phi_{DNA} = 32°$ has $\vec{D}_{Ff(mutant)} = (-45/10/3)$ (*B* of Fig. 2); the minimum at $\Delta\phi_{DNA} = 40°$ has $\vec{D}_{Ff(mutant)} = (-45/10/4)$ (*C* of Fig. 2).

We can now calculate the interaction energy for the $\vec{D}_{fd(native)} = (-60/10/5)$ double resonance structure, as a function of the protein rotation angle. This curve shows a single minimum at $\Delta\phi_{pro} = 36°$, but it is weak because the 16 Å value for Δz_{pro} is very large, in accord with the remarks in the last section. The positive self-energy of the protein coat also shows a small minimum at this angle, because of its tendency to maximally space the repelling charges of its pentamers. Thus, it appears that the 36° Class I protein lattice rotation angle can only be weakly stabilized by an optimized interaction.

Fig. 3 is product lattice for the native Ff model. The fixed DNA strand and the protein lattice both lie on the product lattice, and the second DNA strand can be located at offsets of $(36j)°$ from the fixed strand. In terms of the DNA-protein interaction alone, each of these positions is equivalent, but DNA inter-strand repulsion and stearic constraints would tend to select an azimuthal offset that separates the strands well, such as 180° or 144°. The spacing between points of the product lattice can easily accommodate the four charges of the C-terminus of the Ff coat protein; the adjacent points are neither so far apart that the charges cannot reach them nor so close together that they cannot be distinguished by finite orbitals.

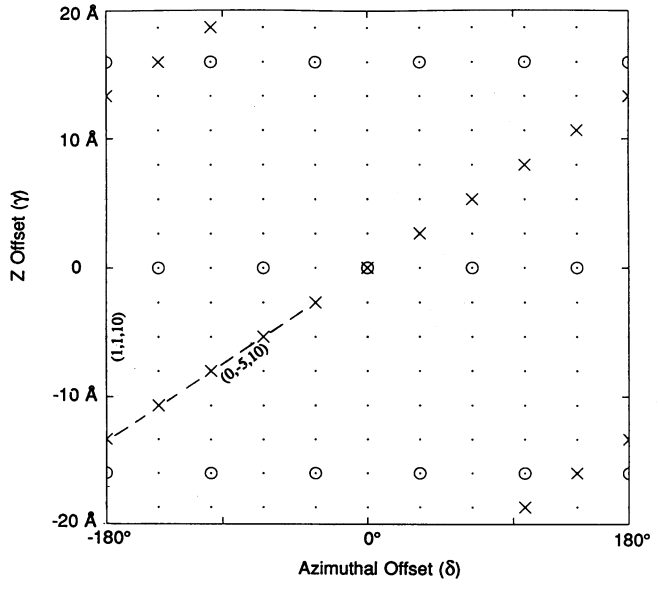

FIGURE 3 A product lattice for the $\vec{D}_{Ff(native)} = (-60/10/5)$ model. The Class I protein lattice (○) and one DNA strand (X) are shown, with one protein charge and one DNA charge at the origin; the abscissa is the azimuthal offset of the second DNA strand, and the ordinate is its *z* offset. The shared resonances (1,1,10) and (0,−5,10) are shown; examples of subhelices that are not resonances are $P^{(pro)}_{0,5}$ and $P^{(pro)}_{1,15}$, shown in Fig. 1 *A*.

Pf1

In the following, we will consider Class II protein coats, single start helices with idealized symmetry parameters $\Delta\phi_{pro} = 66.66°$ and $\Delta z_{pro} = 2.9$ Å, lattice parameters for the low temperature form of Pf1. Fig. 4 shows a large-scale contour plot of DNA-protein interaction energy based on a single idealized Class II protein helix and a single DNA helix. The figure is dominated by a "main sequence," which contains by far the most powerful interactions, a chasm on the free energy contour plot. It represents those DNA helix parameters that correspond to a DNA pitch P_{DNA} equal to the basic protein pitch P_{pro}, which implies $n/s = 2\Delta\phi_{pro}/\Delta\phi_{DNA}$. For structures on the main sequence, the basic protein and DNA helices can be taken as one pair of shared resonant subhelices. Then $P_{DNA} = P_{pro}$ implies $N\,\Delta\phi_{pro}/\Delta z_{pro} = N\Delta\phi_{DNA}/\Delta z_{DNA}$, or $k^{(pro)}_{0,N} = k^{(dna)}_{0,N}$; this resonance may be written $\vec{R} = (0, 0, N)$. For a main sequence double resonance $\vec{D} = (0, 0, N_1)$ X (K_2, M_2, N_2), the corresponding double resonance vector is $\vec{D} = (-M_2/K_2/0)$, so $\vec{D}$ lies in the *KM* plane and its direction is independent of N_1 and N_2. Fig. 5 plots the DNA-protein interaction energy vs. n/s along the main sequence, displaying these troughs in finer detail. Equation 16 says that $n/s = 2(M_2/K_2)$, and minima at these rational numbers are clearly evident, particularly those with $n/s = 1$ and $n/s = 2$.

We now consider Pf1. Its stoichiometry $n/s = 1$ (Kostrikis et al., 1994) immediately suggests that its DNA and protein arrangement might be simple, both helices following the same pitch, with $\Delta z_{DNA} = 2\Delta z_{pro}$, corresponding to the $n/s = 1$ resonance in Fig. 5. Whatever the symmetry of Pf1 DNA, its axial rise of as much as 6.1 Å, the largest of which we are aware, guarantees that it possesses an unusual configuration. This conclusion is supported by the circular dichroism and ultraviolet absorbance spectra of Pf1, which are unlike those of any classical DNA (Casadevall and Day, 1983; Kostrikis et al., 1994). Because the interphosphate distance along a maximally stretched DNA strand is under ~7.5

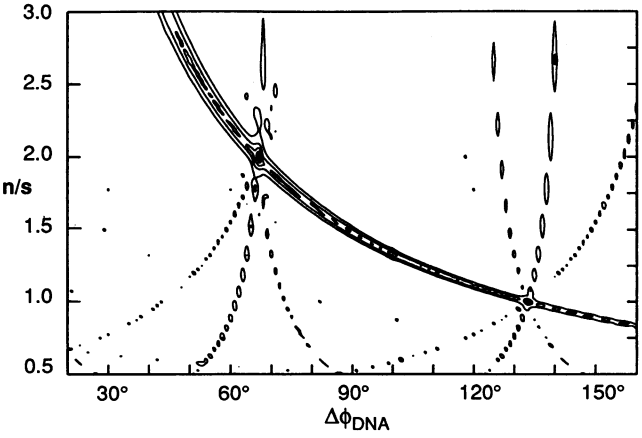

FIGURE 4 A contour plot of interaction energy as a function of $\Delta\phi_{DNA}$ (*x* axis) and n/s (*y* axis) for a Class II protein coat and different single DNA helices coaxial to it. Parameters used for the four-zone model were those used for Fig. 2. The descending arc is the "main sequence," which gives the most powerful resonances possible for any set of input parameters.

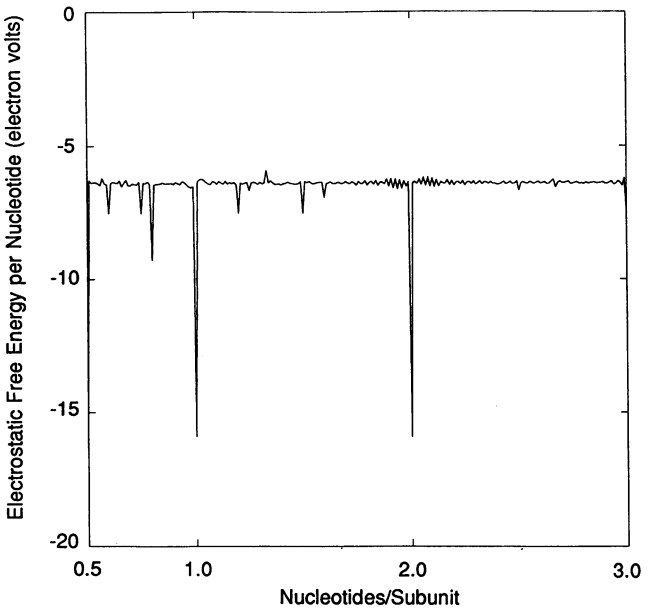

FIGURE 5 Free energy of interaction per nucleotide as a function of n/s for coaxial helices along the main sequence shown in Fig. 4. Along the main sequence the DNA and protein helices have the same pitch and $\Delta\phi_{DNA} = 2\Delta\phi_{pro}/(n/s)$. Each dip is a double resonance, the strongest corresponding to n/s values that are the ratios of small integers. The interaction energy was calculated for 100 DNA residues, with $\kappa = 0.5$ Å^{-1}, $\epsilon = 3$, the DNA charges at 10 Å radius, and the protein charges at 11 Å radius.

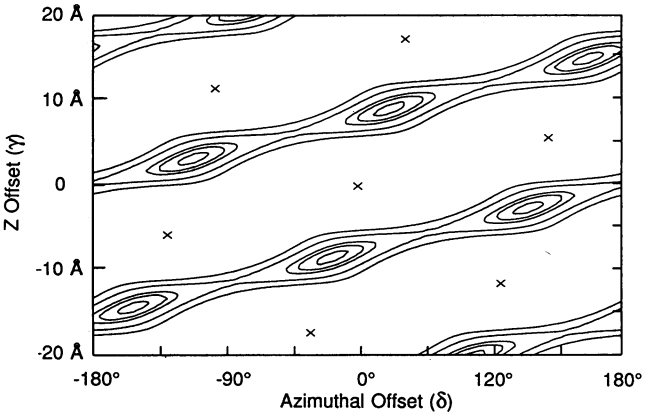

FIGURE 6 A contour plot of total electrostatic free energy per unit length (eV/Å) caused by DNA-DNA interactions; free energy was calculated for a simplified four zone model with $\kappa = 0.2$ Å^{-1}, $\epsilon = 3$, DNA charges at 3.4 Å radius, and protein charges at 7 Å radius. Lattice parameters were (½)-$\Delta z_{DNA} = \Delta z_{pro} = 2.9$ Å and (½)$\Delta\phi_{DNA} = \Delta\phi_{pro} = 65.4545°$. DNA strand 1, denoted by ×, was fixed, with a charge at the origin. The abscissa is the azimuthal offset of strand 2, γ, the ordinate is its z offset, δ. Only the low energy contours are plotted; the contours of increasingly higher energy, at lower offsets from the fixed DNA strand, have been suppressed for clarity. The displacement of the second strand producing the lowest energy has $\gamma = (\Delta\phi_{pro} + \pi)$ and $\delta = \Delta z_{pro}$ (see also Fig. 7).

Å, the backbone of a helical DNA molecule with a given Δz_{DNA} and a given P_{DNA} can be pulled out only to a maximum radius determined by the geometry (Day et al., 1979). Given its extremely large Δz_{DNA} value, a Pf1 DNA that follows the protein helix, and thus has $P_{DNA} \sim 16$ Å, must have its backbone wound tightly around the structure axis; the radial position of the phosphorus atoms is under 2.5 Å. Not until the DNA pitch exceeds about 80 Å can the DNA backbone reach out as far as 10 Å. It is surprisingly easy to make a stereochemically feasible double-stranded DNA molecule with a pitch of 16 Å and an axial rise of as much as 6.1 Å (Day et al., 1988; and our unpublished modeling experiments).

However, if the Pf1 DNA has its sugar phosphate backbone near the center, then the electrostatic repulsion between the two strands assumes paramount significance. Given such a constraint, the structure must be optimized with respect to the electrostatic energy. The most significant constraint involves the phasing of the two strands of DNA, determined by the need to maximize the inter-phosphate distances. Fig. 6 shows contours of the electrostatic free energy of the two DNA strands, as it depends on their relative position. Strand 1 is fixed, denoted by a cross, with one phosphate charge positioned at the origin; the abscissa is the azimuthal offset of strand 2, and the ordinate is its z offset. The entire second strand moves according to these offsets. Because the self energy of one strand is constant, the contours show only the repulsion between the two strands. The minima occur for offsets that locate strand 2 in the potential trough between the windings of strand 1, and within that trough there are two special positions. One is a minimum lying midway between two points of the fixed DNA lattice, on the $\vec{R} = (2, 1, 2)$ resonance (i.e., along the $P_{1,2}^{(pro)} = P_{2,2}^{(DNA)}$ subhelices); the other is a "neck" that lies midway between two other points of the fixed lattice, but on the $\vec{R} = (2, 1, 4)$ resonance. It is clear from inspection that the phosphate charges of the two strands are maximally spaced when a phosphate of strand 2 lies on the minimum.

These special positions can be determined analytically. The electrostatic free energy of interaction of the two DNA point charge helices depends on the relative phasing of the two strands through an exponential factor like that of the s sum of Eq. 18, but written for $k_{kn}^{(DNA)}$ instead of $k_{mn}^{(pro)}$, and summed over indices k and n. Because both strands have negative charges, the s sum vanishes for a given (k,n) if $n\gamma - k_{kn}^{(DNA)}\,\delta = (2L + 1)\pi$; the extra π term on the right-hand side causes cancellation with the contribution from strand 1. Because we are not now considering a resonance with a protein sheath, which would select specific K and N values, we minimize the sum in Eq. 5 by seeking offsets which allow such cancellation for the largest number of (k, n) pairs. Setting $\gamma = a\Delta\phi_{DNA} + c\pi$ and $\delta = b\Delta z_{DNA}$, the equation above for γ and δ shows that the s sum can vanish for an arbitrary $\Delta\phi_{DNA}$ only if $a = b$. Cancellation then demands that $2L + 1 = 2ka + nc$, which is possible for a large set of ks and ns only if c and $2a$ are integers. Because of azimuthal 2π periodicity, only $c = 0$ and $c = 1$ need be considered. Because $a = b$, the possibility $c = 0$ gives strand 2 offset along strand 1, which is surely a maximum overlap, so we have $c = 1$. Because of the helical periodicity along a strand, it is sufficient to consider the cases A) $2a = 0$ and B) $2a = 1$. Case A gives $\gamma = \pi$ and $\delta = 0$, and cancellation occurs if $2L +$

$1 = n$. Inspection of Fig. 6 shows that case A corresponds to the "neck," which could become a minimum if the protein rotation angle were different. Case B has $\gamma = (\frac{1}{2})\Delta\phi_{DNA} + \pi$ and $\delta = (\frac{1}{2})\Delta z_{DNA}$, which gives cancellation if $2L + 1 = k + n$. Case B corresponds to the position of the minimum of Fig. 6. Either case gives no contribution to the free energy from one set of k and n values and maximum contribution from the remaining values. Which of the two minima is deeper therefore depends on which k and n values give the largest A_{kn} values. The Appendix shows how the A_{kn} depend on k, n, and Δz_{DNA}, through $k_{kn}^{(DNA)}$. In a sequence of DNA lattices made with a different $\Delta\phi_{DNA}$ values, the relative strengths of the two minima would change.

The optimal phasing of the DNA helices has consequences for the DNA contribution to the Pf1 fiber diffraction pattern. The Pf1 protein helix parameters (Table 1), $\Delta z_{pro} = 2.77\ A$ and $\Delta\phi_{pro} = 66.66°$ give the selection rule: $l/c = (5n - 27k)/75A^{-1}$; the DNA has the same pitch but twice the axial rise, with selection rule $l/c = (10n - 27k)/150A^{-1}$. (We use index k instead of the usual m to connect consistently with the previous development.) Thus, the DNA layer lines with odd k values correspond to "half layer lines" when indexed on the protein pattern, and the even k lines superimpose onto the protein pattern. For simplicity, the structure factor caused by the phosphate groups, before cylindrical averaging, can be approximated as (Holmes and Blow, 1966):

$$F_{phosphate}(R,\theta) \sim \sum_{m=-\infty}^{m=+\infty}\sum_{k=-\infty}^{k=+\infty} J_n(2\pi R r_{phosphate}) e^{in(\theta+\kappa/2)} \sum_{s=1}^{s=2} e^{-i(n\gamma_s - k_{kn}^{(DNA)}\delta_s)}, \quad (20)$$

where the s sum is over strands 1 and 2. The sum over s is what gave cancellation, for appropriate k and n values, in cases A and B above. Thus, we find immediately for the more likely phasing of the DNA strands of case B, that the phosphates make no contribution to the diffraction when $k + n$ is an odd integer, and so the half layer lines with odd k have no contributions from Bessel functions with even n. For case A the s summation cancels if n is odd for any k, so the half layer lines have no diffraction spots from odd order Bessel functions.

Having considered the optimal theoretical DNA-DNA interaction, we can now consider the protein charges. Fig. 7 is a product lattice diagram for the $\tilde{D}_{Pf1} = (-1/2/0)$ double resonance, showing the optimal phasing positions for the protein and the two DNA helices. Built from the smallest integers that can give rise to $n/s = 1$, this double resonance is very strong, and it also is the simplest that can accommodate DNA strand 1 (X), DNA strand 2 (+), and the two positive charges at the DNA binding C-terminus of the subunit (denoted by a *circle* and a *square*). To minimize the inter-strand repulsion, the charges of the second DNA strand (+) are positioned in the minima of Fig. 6, discussed above. The arrangement of protein charges allows each DNA strand equal access to a protein charge, although it has the X strand always associated with a "circle" charge, and the + strand always associated with a "square" charge. This

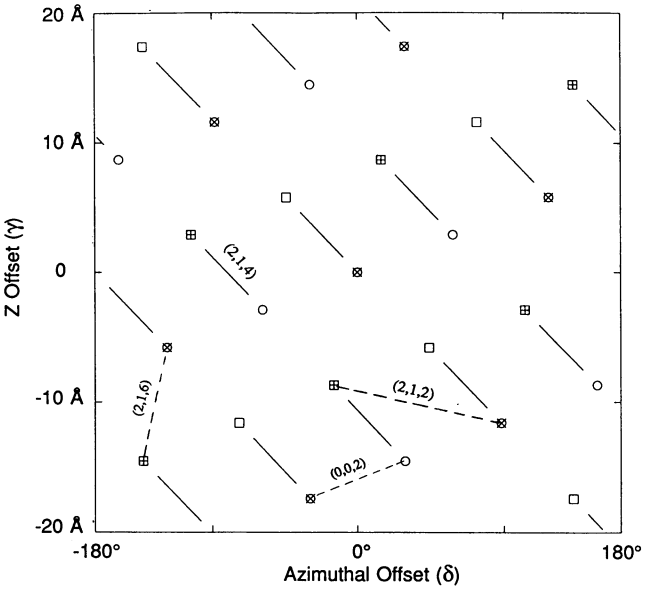

FIGURE 7 A product lattice diagram for the Pf1 model, showing the positions of the charges on the two DNA strands (x,+) and the positions of two charges on each protein subunit (□, ○). This product lattice corresponds to double resonance $\tilde{D} = (-1/2/0)$, with basis vectors (0, 0, 2) and (2, 1, 2); both of these subhelix directions are shown in dotted lines. The negative charges of the first DNA strand are denoted by xs. The second DNA strand is displaced from the first along the (2, 1, 2) direction by offsets $\gamma = (\Delta\phi_{pro} + \pi)$ and $\delta = \Delta z_{pro}$, and its negative charges are denoted by cross-marks (+). The two positive charges associated with a single protein subunit (□, ○) are displaced from each other along the (2, 1, 4) direction with offsets $\alpha = (2\Delta\phi z_{pro} + \pi)$ and $\beta = 2\Delta z_{pro}$; these charges are linked by solid lines. Unlike other product lattice diagrams shown below, there are no unoccupied lattice points.

lattice has an azimuthal displacement of 47° between protein charges from one subunit, and their Z spacing is about 6 Å; they are displaced along the $P_{1,4}^{(pro)}$ subhelix direction, the subhelix of the $\tilde{R} = (2, 1, 4)$ resonance. With the axis of the C-terminal helix axis at a radius of about 15 Å, the required spacings can be attained with the long side chains of Arg^{44} and Lys^{45}, maximally splayed. The strong local potential fields help to position these charges.

In the considerations leading to Fig. 7, the idealized protein rotation angle $\Delta\phi_{pro} \sim 66.66°$ was assumed. This assumption can be explored by varying $\Delta\phi_{pro}$, while maintaining the value $n/s = 1$, the equality of the DNA and protein pitches (main sequence), and the resonant phasing relationships between the DNA strands and the protein charges of Fig. 7. The resulting sequence of structures is what one would make by redrawing this figure, using a range of values for $\Delta\phi_{pro}$, but maintaining the relationship $\Delta\phi_{DNA} = 2\Delta\phi_{pro}$ and all z values. The total electrostatic free energy of this model includes the DNA-protein interaction energy and the DNA-DNA and protein-protein self-energies, calculated according to the phasings used in Fig. 7; thus, it includes the two strands of DNA and two charged helices generated by the doubly-charged protein. Fig. 8 shows that A/L has a minimum for $\Delta\phi_{pro} \sim 65°$, which is very near the observed value of $\sim 66.6°$. The minimum occurs at somewhat lower protein

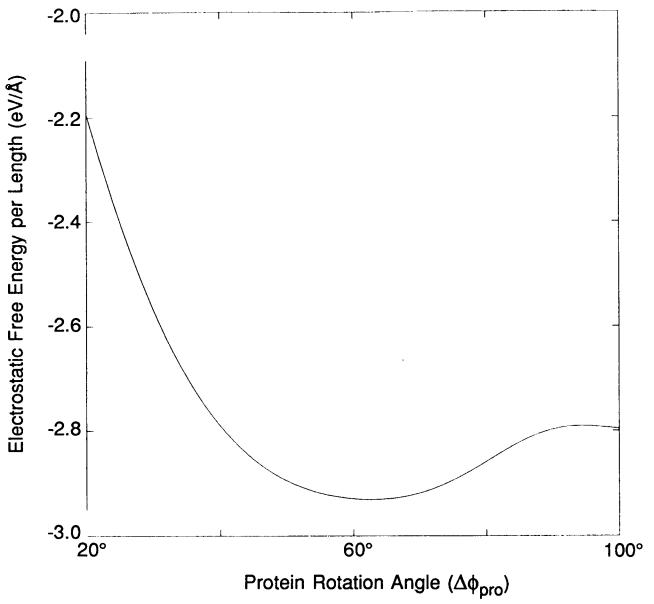

FIGURE 8 Electrostatic interaction free energy per unit length along structure axis, expressed as electron volts per Å, as a function of protein rotation angle $\Delta\phi_{pro}$ for the Pfl model. The interaction energies were calculated on the basis of two DNA strands and a doubly-charged protein subunit, with relative phasings dependent on $\Delta\phi_{pro}$, as depicted in Fig. 7. For this calculation, $\kappa = 0.3$ Å^{-1}, $\epsilon = 3$, the DNA charges were at 2.75 Å radius, and the protein charges were at 5 Å radius. Bound models occur if the protein charge is moved as far out as 8 Å radius.

rotation angles, usually between 60° and 64°, for different choices of modeling parameters. Those parameter sets that give the most negative free energy also give $\Delta\phi_{pro}$ values closest to 66.66°; when the neutralizing protein charges are brought to between 1.5 and 2 A of the phosphate charges, the smallest free energy occurs at ~66.6°.

This concurrence between the theory and the empirical rotation angle of the protein lattice depends intimately on the presence and phasing of the two DNA strands and the two charged protein helices. With its phosphate atoms lying within ~2.5 Å from the structure axis, the DNA model has phosphate charges within ~4 A from the structure axis. To maintain an electrostatically bound structure, one with a negative total free energy, the "neutralizing" protein charges must lie within ~8 Å from the structure axis. They can do this because they are situated on the long side chains of lysine and arginine residues, at the C-terminus of the subunit. However, at such a small radius and separated by only one-half of the rotation angle and axial rise of the DNA, these charges confer a substantial positive, de-stabilizing, self-energy to the protein helix. To see the effects of the various contributions to the total electrostatic free energy, we can remove them one at a time from the full model. Removing all energies except the interaction energy between a single strand of DNA and a single protein helix gives a free energy minimum at about $\Delta\phi_{pro}$ ~100°. Removing only the protein self-interaction gives no minimum in A/L in the range 20° < $\Delta\phi_{pro}$ < 120°. If only one of the two protein charges is allowed to approach the DNA charges (i.e., if either the *circles* or the *squares* are removed from Fig. 7), the minimum is near 80° for parameters that give a strongly bound structure, and it approaches 65° only for parameters that give very weak binding. Thus, it appears that both protein charges must be included to find a minimum of A/L at a protein rotation angle close to the observed value for very stable structures. This particular result is model-dependent, but we believe that the model proposed already captures the most important features. By contrast, the argument giving the relative phasing of the two DNA strands is geometrical, and so it is not model-dependent.

Marvin et al. (1992) have sketched a Pf1 model in which a double resonance appears implicitly. To avoid the electrostatic problem attached to phosphates-in DNA, they locate the phosphorus atoms at a radius of about 5 Å. Their proposed DNA pitch is 185 Å, corresponding to the $P_{1,6}^{(pro)}$ subhelix of the protein lattice. Their DNA and protein lattices are in the weak double resonance: $\vec{D} = (3/-6/1)$, with two basis resonance vectors $\vec{R} = (0,1,6)$ and $\vec{R} = (-1, 1, 9)$; this double resonance has a deep minimum in its interaction free energy for a protein rotation angle of 60°, so it does not stabilize the observed $\Delta\phi_{pro}$ of 66.6°.

Xf

Early studies of Xf suggested that its n/s value was the integer 2. However, more recent results now indicate a noninteger value greater than 2 (Table 1), and silver binding studies (Casadevall and Day, 1982, 1983) show that Xf DNA is right-handed with a pitch in the vicinity of 30 Å, closely similar to Ff DNA. Right-handedness means that only the positive rotation angle part of the n/s ~$\Delta\phi_{DNA}$ plane in the range of 25°–40° need be considered. Because this range excludes the main sequence, its resonances are less powerful than that of Pf1. Allowing for error in the n/s values, we consider only models with $2.1 < n/s < 2.3$. Within this range, six solutions appear possible, and the LPB model estimates that their strengths, compared with the off-resonance background, are similar. The strongest double resonance of these is $\vec{D} = (12/-11/1)$, with $n/s = 2(12/11) = 2.1818$, $\Delta\phi_{DNA} = 31.111°$, and $P_{DNA} = 30.7$ Å from Eqs. 14 and 16. These parameters agree well with observations, but this double resonance has its minimum free energy at a protein rotation angle of ~72°, so this model for the Xf virion is not an optimized double resonance. Two optimized double resonances occur within the allowed range: the first is $\vec{D} = (13/-12/1)$, with $n/s = 2(13/12) = 2.1667$, $\Delta\phi_{DNA} = 33.84°$, and $P_{DNA} = 27.3$ Å; The second is $\vec{D} = (37/-35/3)$, with $n/s = 2(37/35) = 2.1143$, $\Delta\phi_{DNA} = 33.87°$, and $P_{DNA} = 27.9$ Å. These DNA pitches and rotation angles are based on a protein rotation angle of 66.66° and an axial rise of 2.77 Å. These three vectors point in neighboring directions in *KMN* space.

The resonant interaction idea cannot discriminate unambiguously between the last two possibilities, but the choice is clarified by comparing the two corresponding product lattices. Fig. 9 shows the $\vec{D} = (13/-12/1)$ product lattice, characterized by a prominent line of points along the direction of

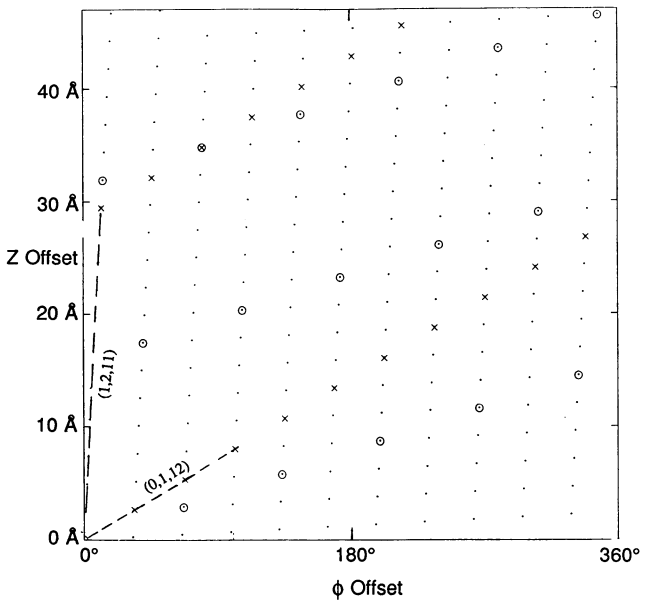

FIGURE 9 The product lattice for the Xf model, the $\vec{D}$ = (13, −12, 1) double resonance; it was calculated using $\Delta\phi_{pro}$ = 2.777 Å and $\Delta\phi_{pro}$ = 66.666°. Directions (1, 2, 11) and (0, 1, 12) are indicated; the (1, 2, 11) direction is nearly parallel to the structure axis because this double resonance is nearly optimum; i.e., 66.66° ~ (2/11)360°. Circles represent one protein charge of each subunit, and xs represent one DNA strand. The other three protein charges and the other DNA strand would generate helices parallel to the protein and DNA basic helices, respectively; they would also occupy the product lattice.

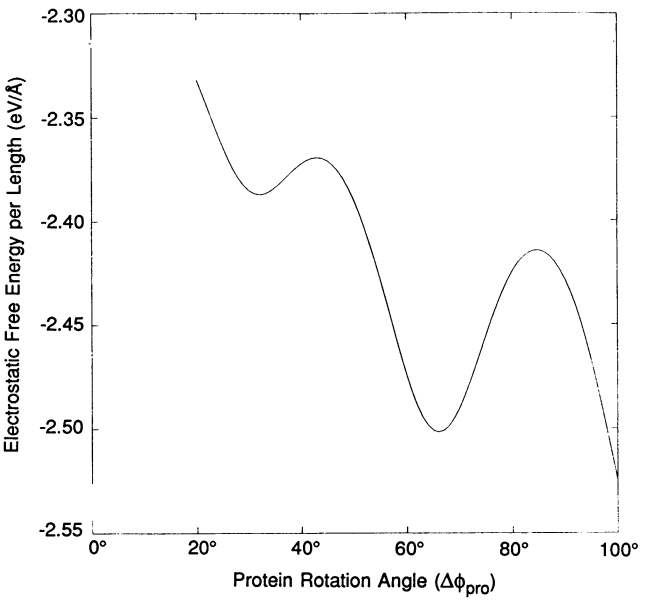

FIGURE 10 Electrostatic interaction free energy per unit length (eV/Å) vs. protein rotation angle, $\Delta\phi_{pro}$, for the Xf double resonance shown in Fig. 9. It shows minima at about (2/11) 360° and (1/12) 360°; these dominate the figure because the smallest N values of this double resonance family are 1, 11, and 12.

the (1, 2, 11) resonance, a member of the $\vec{D}$ = (13/−12/1) family. The $P_{2,11}^{(pro)}$ subhelix has appeared prominently in the modeling of the Class II protein lattice (Makowski et al., 1980; Makowski and Caspar, 1981; Marzec and Day, 1983, 1988), as the direction of the protein subunit backbone in the DNA interaction domain. In the product lattice, there are 13 points between successive subunits along the (1, 2, 11) subhelix direction, each spaced in the z direction by 11(2.77)/13 Å = 2.35 Å. This configuration makes it particularly easy for the four positive charges of the C-terminal region of Xf to lie on the well spaced product lattice. The points of the (37/−35/3) product lattice also follow the subunit backbone, but 37 points are found between successive subunits, placing them only about 0.8 Å apart in the z direction; as discussed in the Ff section, this would be too close. Thus, for Xf our working hypothesis picks out the $\vec{D}_{Xf}$ = (13/−12/11) double resonance, with n/s = 2.167. Fig. 10 shows the electrostatic interaction free energy per length, A/L, as a function of $\Delta\phi_{pro}$ for this model, with a minimum located near $\Delta\phi_{pro}$ ~ (2/11)360°, as expected. Similar plots for the rejected double resonances within the allowed region show minima near 60° and 72°, which are far from the Class II protein rotation angle.

Pf3

The data available for Pf3 indicate a n/s value near 2.4, a DNA structure that probably is not classical (Casadevall and Day, 1983; Day et al., 1988), and a Class II fiber diffraction pattern (Peterson et al., 1982). We can use the double resonance hypothesis to delineate models that are consistent with these data. Because Pf3 has a Class II fiber diffraction pattern, the optimized double resonance working hypothesis says that one of the most powerful resonances of Pf3 has the form $\vec{R}$ = (K_1, 2, 11), for an unknown integer K_1; the (M, N) = (2,11) values give $\Delta\phi_{pro}$ near 66°. Assuming that $\Delta\phi_{pro}$ = 66.66°, Eqs. 14 and 16 can be used to find possible n/s and $\Delta\phi_{DNA}$ values, by inserting all reasonable values for K_1, and (K_2, M_2, N_2). This results in Fig. 11, which plots the calculated n/s values against the calculated $\Delta\phi_{DNA}$ values, showing in a multi-branched plot indexed by K_1 values of −4 to 4. The negative rotation angles correspond to a DNA with the opposite hand from the protein; the assumed Class II protein lattice is right-handed, and there is no information about the handedness of the DNA. Each branch corresponds to all choices of (K_2, M_2, N_2), for a given K_1. (The points of each branch fall on a one-dimensional curve instead of a two-dimensional scatter plot because every possible double resonance $\vec{D}$ associated with any given $\vec{R}_1$ = resonance vector must be orthogonal to it: $\vec{D} \cdot \vec{R}$ = 0; thus, the double resonances which can include $\vec{R}1$ belong to a one parameter family.)

The possible $\Delta\phi_{DNA}$ values are approximately quantized, having values that are multiples of (2/11) 360°. Possible Pf3 models must have the $\Delta\phi_{DNA}$ values at which the branches intersect the line n/s ~2.4. According to the optimized double resonance hypothesis, any Class II virion corresponds to some point on this figure, with its quantized DNA rotation angles. The Pf1 model corresponds to the n/s = 1 point on the K_1 = 4 branch, and the two Xf models discussed above

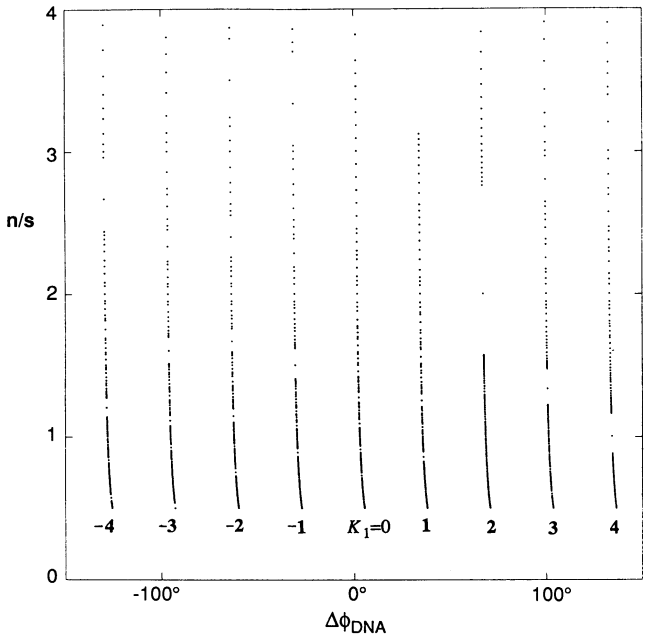

FIGURE 11 n/s vs. $\Delta\phi_{DNA}$, based on an assumed value of $\Delta\phi_{pro} = 66.666°$ and an assumed resonance vector $\vec{R}_1 = (K_1, 2, 11)$. Each point corresponds to a double resonance which includes $\vec{R}_1$ in its plane of resonances, and each double resonance generates an n/s value and a value for $\Delta\phi_{DNA}$. Each K_1 value generates a different $\vec{R}_1$ and a new set of possible double resonances that include $\vec{R}_1$. For $\Delta\phi_{pro} = (2/11)360°$, the curves become straight exactly vertical lines.

fall on the $K_1 = 1$ branch. It is interesting that gaps appear on each branch for certain n/s ranges, forbidden values.

DISCUSSION

We have calculated numerical values for free energies via the LPB equation only to compare double resonances and to ascertain that the models presented are electrostatically bound; e.g., that the electrostatic free energy of Pf1 is negative for physically realistic modeling parameters. Because the LPB equation is inappropriate for electrostatics on such small spatial scales and its input parameters are poorly known anyway, our numerical results are only order of magnitude guidelines. We have found in every case that it is quite easy to make models with 1–10 ev. per nucleotide of binding energy, and on this basis we conclude that the energy difference between a doubly resonant structure and an off-resonance structure is physically significant.

The resonance condition, which says that protein and DNA basic lattices share one subhelix, is a generalization of the pitch connection relation, derived earlier from mechanical considerations; the double resonance is a generalization of the restricted pitch connection relation (Marzec and Day, 1983). In that work, the allowed M and N values were limited to those that characterize the alpha helix tubes of the protein models: Class I viruses had $M/N = 2/10$ and Class II had $M/N = 2/11$. The restrictions on M and N derived from the mechanical means used to relate helices. They do not apply to the electrostatic model for the interaction because the immateriality of the potential field allows for contributions from subhelices with all M and N values.

Those M/N pairs of a double resonance family that have relatively small N values generate minima of the interaction energy for optimized values of $\Delta\phi_{pro} \sim (M/N)\,360°$. We have argued elsewhere (Marzec and Day, 1988) that there exists a set of M/N pairs corresponding to rotation angles that generate filamentous virus models that close pack minimally curved alpha-helices around the DNA core. This concurrence is not accidental, because both minimizations arise from a $\Delta\phi_{pro}$ value that causes a relevant k_{mn} to vanish. Both the packing of α-helices and the interaction of coaxial helices are described in terms of k_{mn} values and their associated subhelices because these concepts belong to the natural language for describing helical lattices.

The sort of analysis that we have given, for regular structures bearing a global symmetry and predicated on long-ranged, nonspecific interactions, allows for small variations in the shapes and positioning of the morphological units from which the structure is built. We believe that this idea applies with some generality. In the case of the filamentous bacteriophages organized into coaxial helices, the morphological unit of the protein helix is a monomeric protein subunit, and that of the DNA helix is a nucleotide. The irregular sequence of bases implies that the nucleotide morphological units are not identical. Also, n/s ratios that do not exactly equal two indicate a DNA-protein interaction that must vary from subunit to subunit. Together, these features must induce small variations in protein structure at the DNA-protein interface, giving protein morphological units that are also not identical in shape. Thus, we anticipate a tendency for the DNA molecule to deform somewhat from a perfect helix, as nucleotides adopt their own lowest energy positions, having DNA-protein interaction somewhat stronger than that given by DNA in a perfect helix. Similarly, the helix of protein subunits will also undergo small variations. The isometric spherical viruses represent another case of a regular structure with a symmetry induced by nonspecific interactions (Marzec and Day, 1993). Their capsid morphological units are typically polymeric capsomeres that, because of the underlying icosahedral symmetry of the capsid, cannot be in identical environments (except for those viruses with triangulation number $T = 1$). In both the filamentous and spherical structures, each morphological unit interacts not only with its immediate neighbors, but with many others, registering the global symmetry that informs the structure and positioning itself accordingly. In both cases, a small variation in the structure or the position of a morphological unit can be thought of as inducing a first-order perturbation in the physical fields carrying the global symmetry; the variations associated with any two morphological units thus compound nonlinearly, to relatively negligible second order. Accordingly, we have set aside consideration of any nonlinear perturbations of the helical symmetry itself, and our analysis herein has been linear, concerned with the

interactions of unperturbed, coaxial helices. Similarly, the linearized form of the Poisson-Boltzmann equation has been employed. It might be of interest to calculate nonlinear contributions to the Helmholtz free energy, to determine whether they are of more significance for some minima than for others, possibly to remove some minima from consideration as solutions for particular virions. However, the *n/s* and rotation angles of the minima would certainly be unchanged by a nonlinear analysis.

The structure of a macromolecular assembly containing a genome must accommodate the natural variability of its nucleotide sequence, and it must be stable to both large and small alterations in that sequence caused by mutations, insertions, and deletions. We have argued that structures determined by repeating, nonspecific, long-ranged interactions are essentially unchanged by local variations in their sub-assemblies (e.g., a filamentous bacteriophage with morphology unaltered by changes in its nucleotide sequence or in its amino acid sequence). Thus, we infer that stable storage of discrete (digital) information is achieved via structures determined by long-ranged, nonspecific interactions.

CONCLUSION

The optimized double resonance hypothesis offers a cogent framework that addresses naturally the symmetry matching problems presented by coaxial helical structures. Existing data can be interpreted; and the hypothesis makes relative and absolute predictions about helix parameters and suggests optimum phasings for interacting charges. However, so few of the coaxially symmetrical structures that might display double resonance symmetry matching have been measured that the hypothesis must await more data for corroboration.

This work supported through National Institutes of Health grant GM 42286-24.

APPENDIX

The general form of a helically symmetrical field

The potential field created by a helical distribution of point charges displays the geometrical symmetry of its sources, and this fact leads to a general form for such a field. In cylindrical coordinates, the field has the Fourier expansion:

$$V(r,\phi,z)=\sum_{m=-\infty}^{m=+\infty}\sum_{n=-\infty}^{n=+\infty}A_{mn}(r)\exp\{i[n(\phi-\phi_0)-k_{mn}(z-z_0)]\} \tag{A1}$$

Here, we anticipate that because of the helical symmetry, the wave number k will be a discrete variable, labeled by indices m and n. The helical symmetry requires:

$$V(r,\phi+\Delta\phi,z+\Delta z)=V(r,\phi,z) \tag{A2}$$

for all z and ϕ. The variables Δz and $\Delta\phi$ are the axial rise and rotation angle of the basic helix, with pitch $P=2\pi\Delta z/\Delta\phi$. Evaluated at $\phi+\Delta\phi$ and $z+\Delta z$, (A1) gives

$$V(r,\phi+\Delta\phi,z+\Delta z) \tag{A3}$$
$$=\sum_{m=-\infty}^{m=+\infty}\sum_{n=-\infty}^{n=+\infty}A_{mn}(r)\exp\{i[n(\phi-\phi_0)-k_{mn}(z-z_0)]\}\exp\{i(n\Delta\phi-k_{mn}\Delta z)\}.$$

The right-hand sides of Eq. A1 and A3 are equal if and only if

$$k_{mn}=\frac{n\Delta\phi-2\pi m}{\Delta z} \tag{A4}$$

for m an integer. We can define a pitch

$$P_{mn}\equiv\frac{2\pi n}{k_{mn}}=\frac{2\pi n\Delta z}{n\Delta\phi-2\pi m}. \tag{A5}$$

Thus, we have the equivalent forms for the helically symmetrical field:

$$V(r,\phi,z)=\sum_{m=-\infty}^{m=+\infty}\sum_{n=-\infty}^{n=+\infty}A_{mn}(r)\exp\{in[(\phi-\phi_0)-2\pi(z-z_0)/P_{mn}]\}$$

$$=\sum_{m=-\infty}^{m=+\infty}\sum_{n=-\infty}^{n=+\infty}A_{mn}(r)\exp[in(\phi-\phi_0)] \tag{A6}$$
$$\times\exp[i(2\pi m-n\Delta\phi)(z-z_0)/\Delta z]$$

$$=\sum_{m=-\infty}^{m=+\infty}\sum_{n=-\infty}^{n=+\infty}A_{mn}(r)\exp[in(\phi-\phi_0)]\exp[-ik_{mn}(z-z_0)]$$

The first form of Eq. A6 shows that $A_{mn}(r)$ may be thought of as the "strength" of a subhelix with pitch P_{mn}. The geometrical interpretation of these pitches is shown in Fig. 1 *a*. It is the province of physics to evaluate the $A_{mn}(r)$ for any given problem.

For the case of an N_R-fold rotationally symmetrical system, we apply the second constraint

$$V\left(r,\frac{\phi+2\pi}{N_R},z\right)=V(r,\phi,z). \tag{A7}$$

Evaluated at $(r,\phi+2\pi/N_R,z)$, (A1) gives

$$V\left(r,\frac{\phi+2\pi}{N_R},z\right) \tag{A8}$$
$$=\sum_{m=-\infty}^{m=+\infty}\sum_{n=-\infty}^{n=+\infty}A_{mn}(r)\exp\left\{i\left[n\left(\phi-\phi_0+\frac{2\pi}{N_R}\right)-k_{mn}(z-z_0)\right]\right\}.$$

The right-hand sides of Eqs. A1 and A8 are equal if and only if

$$A_{mn}(r)=0 \quad \text{unless} \quad n=sN_R \tag{A9}$$

for some integer s; i.e., n must be an integral multiple of N_R. All equations in this Appendix can be generalized to the rotational case by replacing sums over index n with sums over index s and then setting $n=sN_R$. So the rotationally symmetrical form of Eq. A6 is

$$V(r,\phi,z)=\sum_{m=-\infty}^{m=+\infty}\sum_{s=-\infty}^{s=+\infty}A_{m,sN_R}\exp\left\{isN_R\left[(\phi-\phi_0)-\frac{2\pi(z-z_0)}{P_{m,sN_R}}\right]\right\} \tag{A10}$$

where

$$P_{m,sN_R}\equiv\frac{2\pi sN_RT}{sN_R\Delta\phi-2\pi m}=\frac{2\pi sN_R}{k_{msN_R}}. \tag{A11}$$

We have used $T=\Delta z$, the z spacing between N_R-mers, and explicity indicated that n must be a multiple of N_R, but the helical and rotational forms of P_{mn} and k_{mn} are identical. It will be useful below to display the form of the charge density created by an infinite helical distribution of point charges of magnitude Q. The charges have cylindrical coordinates $(R,\phi_0+n\Delta\phi,z_0+n\Delta z)$, where the integer n ranges from $-\infty$ to $+\infty$. This

discontinuous helix can be thought of as a continuous helix, defined by the equation $\phi - \phi_0 = 2\pi(z - z_0)/P$, combined with the requirement that the z coordinates of the charges be distributed discretely, given by $z - z_0 = n\Delta z$. Thus, the charge distribution can be written via Dirac delta functions:

$$\rho(r, \phi, z) = Q\frac{\delta(r - R)}{r}\sum_{n=-\infty}^{n=+\infty}\delta\left[\phi - \phi_0 - \frac{2\pi(z - z_0)}{P} - 2\pi n\right] \times \sum_{m=-\infty}^{m=+\infty}\delta(z - z_0 - m\Delta z) \quad \text{(A12)}$$

When this expression is integrated over a volume element $d^3r = r\,dr\,d\phi\,\Delta z$ that contains a charge, the result must be Q. The first delta function of Eq. A12 requires that the r integration include radius R. The second contains the expression $\phi - 2\pi z/P$, which ensures that the charge density follows a continuous helix of pitch P; the term $-2\pi m$ allows for the fact that the charge density must be periodic in azimuthal angle ϕ. The third delta function breaks the continuous helix into a discontinuous helix, with charges positioned at z values of $m\Delta z$.

Eq. A12 can be rewritten into a more useful form by means of the identities

$$\sum_{n=-\infty}^{n=+\infty}\delta(kx - 2\pi n) = \frac{1}{2\pi}\sum_{n=-\infty}^{n=+\infty}e^{inkx} \quad \text{and} \quad \text{(A13)}$$

$$\sum_{n=-\infty}^{n=+\infty}\delta\left(k - \frac{2\pi n}{x}\right) = \frac{x}{2\pi}\sum_{n=-\infty}^{n=+\infty}e^{inkx}.$$

Letting $kx = \phi - \phi_0 - 2\pi(z - z_0)/P$ in the first δ function of Eq. A13, we find for the second delta function of Eq. A12

$$\sum_{n=-\infty}^{n=+\infty}\delta\left[\phi - \phi_0 - \frac{2\pi(z - z_0)}{P} - 2\pi n\right] = \frac{1}{2\pi}\sum_{n=-\infty}^{n=+\infty}\exp\left\{in\left[\phi - \phi_0 - \frac{2\pi(z - z_0)}{P}\right]\right\}. \quad \text{(A14)}$$

Applying the second of Eq. A13 to the third delta function of Eq. A12, letting $k = z - z_0$ and $x = 2\pi/\Delta z$, we find

$$\sum_{m=-\infty}^{m=+\infty}\delta(z - z_0 - m\Delta z) = \frac{1}{\Delta z}\sum_{m=-\infty}^{m=+\infty}\exp\left[\frac{2\pi im(z - z_0)}{\Delta z}\right]. \quad \text{(A15)}$$

Thus, the density can be written

$$\rho(r, \phi, z) \quad \text{(A16)}$$
$$= Q\frac{\delta(r - R)}{2\pi r\Delta z}\sum_{m=-\infty}^{m=+\infty}\sum_{n=-\infty}^{n=+\infty}\exp[in(\phi - \phi_0)]\exp\left[2\pi i\left(\frac{m}{\Delta z} - \frac{n}{P}\right)(z - z_0)\right].$$

Using the definition of P, it is apparent that Eq. A16 has the same form as Eq. A6, with $A_{mn}(r) = Q\delta(r - R)/(2\pi r\Delta z)$; thus, the $A_{mn}(r)$ distribution that describes the charge density is independent of m and n for a helix of point charges. If the charge density were N_R-fold rotationally symmetrical, then again we would require $A_{mn}(r) = 0$ unless n is an integral multiple of N_R, in which case $A_{mn}(r) = N_R Q\,(\delta(r - R)/2\pi r\Delta z)$, the factor of N_R inserted to account for the N_R-fold increase in the charge per length.

The electrostatic potential of a helix of point charges

We begin with the LPB equation for a cylindrically symmetrical potential $V(r)$ written in cylindrical coordinates:

$$\frac{1}{r}\frac{d}{dr}\left[\epsilon(r)r\left(\frac{dV}{dr}\right)\right] - \kappa^2(r)V = -4\pi\rho_{\text{fixed}}. \quad \text{(A17)}$$

(The "fixed" charges are the source charges of the potential distribution, as distinguished from the counterions.) We will first suppose that the charge density is uniformly distributed on the surfaces of three concentric cylinders:

$$\rho_{\text{fixed}} \quad \text{(A18)}$$
$$= \left(\frac{1}{2\pi r}\right)[\lambda_{\text{DNA}}\,\delta(r - r_{\text{DNA}}) + \lambda_{\text{protein}}\,\delta(r - r_{\text{protein}}) + \lambda_{\text{surface}}\,\delta(r - r_{\text{surface}})].$$

Here the λs are the linear charge densities of the three cylinders, and $\delta(x)$ represents the Dirac delta function. The three cylinders divide the space into four regions. We will permit κ and ϵ to possess different values in each region. This formulation is intended to represent a simple electrostatic model of the virions, allowing us to consider effects resulting from only the radial part of the DNA, protein, and surface charge distributions. In the virion, these charges are actually helically distributed, which presents complications elaborated below.

Within region i, κ and ϵ are constant, assuming values κ_i and ϵ_i, giving

$$\frac{1}{r}\frac{d}{dr}\left[r\left(\frac{dV_i}{dr}\right)\right] - \left[\frac{\kappa_i^2}{\epsilon_i}\right]V_i = 0. \quad \text{(A19)}$$

If the κ^2 term had a + sign, this would be the equation for Bessel functions of zero order. As it stands, Eq. A20 is the equation for modified Bessel functions I_0 and K_0 (Abramowitz and Stegun, 1964) (where Bessel function K_n is not to be confused with index K used in the text), with the solution

$$V_i(r) = a^i I_0(\kappa_i r/\sqrt{\epsilon_i}) + b^i K_0(\kappa_i r/\sqrt{\epsilon_i}) \quad \text{(A20)}$$

in region i (= I, II, III, or IV). The modified Bessel functions I_n and K_n are calculated numerically, with the 0th and first orders found by series expansions, and the higher orders, needed below, by Miller's method for stable recursion (Abramowitz and Stegun, 1964).

The solution of the problem now requires determining the eight coefficients a^i and b^i. To keep $V(r)$ finite along the structure axis, we need $b^1 = 0$, because K_0 diverges as its argument vanishes. Similarly, we need $a^4 = 0$ to keep the potential finite at infinity. Therefore, six more boundary conditions are needed.

It is necessary that $V(r)$ be continuous across each of the three boundaries of the region, as usual, to avoid infinite forces at the boundaries, yielding three boundary conditions. The remaining three boundary conditions are found by multiplying Eq. A17 by $r dr$ and integrating across each of the three boundaries, giving for the I-II boundary:

$$\epsilon_2\left[\frac{dV_2}{dr}\right]\Bigg|_{\text{DNA}} - \epsilon_1\left[\frac{dV_1}{dr}\right]\Bigg|_{\text{DNA}} = -\frac{2\lambda_{\text{DNA}}}{r_{\text{DNA}}}. \quad \text{(A21)}$$

The II-III and III-IV boundaries give similar equations, for a total of three. The resulting system of six equations for the six remaining as and bs has the form $AX = B$, where x is a vector made of the as and bs, A is a 6 by 6 matrix, and B is a vector made from the λs. This system is easily inverted via a standard linear equations subroutine package.

We will now consider the case of a charged helix that lies on one of the boundary cylinders that separate the space into four regions, as above, so $\epsilon = \epsilon(r)$ and $\kappa = \kappa(r)$. The potential field created by a single discontinuous helix can be calculated from the LPB equation in cylindrical coordinates:

$$\vec{\nabla}\cdot[\epsilon\vec{\nabla}V(r, \phi, z)] = \frac{1}{r}\frac{\partial}{\partial r}\left(\epsilon r\frac{\partial V}{\partial r}\right) + \frac{\epsilon}{r^2}\frac{\partial^2 V}{\partial\phi^2} + \epsilon\frac{\partial^2 V}{\partial z^2} = -4\pi\rho_{\text{fixed}} + \kappa^2 V \quad \text{(A22)}$$

The potential V must have the same helical symmetry as the charge density ρ_{fixed} that generates it, and we assume here that ρ_{fixed} possesses

only one helical symmetry. Insert Eq. A1 for V and Eq. A11 for ρ:

$$\sum_{m=-\infty}^{m=+\infty}\sum_{n=-\infty}^{n=+\infty}\exp\{i[n(\phi-\phi_0)-k_{mn}(z-z_0)]\}$$
$$\times\left\{\frac{1}{r}\frac{\partial}{\partial r}\left[\epsilon r\frac{\partial A_{mn}(r)}{\partial r}\right]-\epsilon\frac{n^2}{r^2}A_{mn}(r)-[-\epsilon k_{mn}^2+\kappa^2]A_{mn}(r)\right\} \quad \text{(A23)}$$
$$=-4\pi Q\frac{\delta(r-R)}{2\pi r\,\Delta z}\sum_{m=-\infty}^{m=+\infty}\sum_{n=-\infty}^{n=+\infty}\exp[in(\phi-\phi_0)]$$
$$\times\exp\left[2\pi i\left(\frac{m}{\Delta z}-\frac{n}{P}\right)(z-z_0)\right]$$

The value of k_{mn} is determined by Eq. A4, and Δz and P describe the symmetry of the charged helix. To hold for each ϕ and z, the equality must hold for each pair of indices m and n, so we have

$$\frac{1}{r}\frac{\partial}{\partial r}\left[\epsilon r\frac{\partial A_{mn}(r)}{\partial r}\right]-\epsilon\frac{n^2}{r^2}A_{mn}(r)-[-\epsilon k_{mn}^2+\kappa^2]A_{mn}(r) \quad \text{(A24)}$$
$$=-2\lambda\frac{\delta(r-R)}{r}$$

where $\lambda \equiv (Q/\Delta z)$. Within region i, κ and ϵ are constant, and so we have

$$\frac{1}{r}\frac{\partial}{\partial r}\left[r\frac{\partial A_{mn}(r)}{\partial r}\right]-\frac{n^2}{r^2}A_{mn}(r)-\left[-k_{mn}^2+\frac{\kappa_i^2}{\epsilon_i}\right]A_{mn}(r)=0. \quad \text{(A25)}$$

If $n = m = 0$, Eq. A25 is identical to Eq. A19, which describes the cylindrically symmetrical case, whose solutions are modified Bessel functions of order zero. As written, Eq. A25 is the equation for modified Bessel functions of order n, with solutions given by

$$A_{mn}^i(r)=a_{mn}^i I_n[r\sqrt{k_{mn}^2+\kappa_i^2/\epsilon_i}]+b_{mn}^i K_n[r\sqrt{k_{mn}^2+\kappa_i^2/\epsilon_i}]. \quad \text{(A26)}$$

The coefficients a_{mn}^i and b_{mn}^i are determined as above for the cylindrically symmetrical case. We set $b_{mn}^1 = a_{mn}^4 = 0$, so for each pair m and n, it is necessary to determine six remaining coefficients. Continuity of the potential across the cylinders gives three equations:

$$A_{mn}^1|_{r_{DNA}}=A_{mn}^2|_{r_{DNA}}$$
$$A_{mn}^2|_{r_{protein}}=A_{mn}^3|_{r_{protein}} \quad \text{(A27)}$$
$$A_{mn}^3|_{r_{surface}}=A_{mn}^4|_{r_{surface}}$$

Integration of Equation A-24 across the three boundaries yields three more:

$$\epsilon_2\frac{\partial A_{mn}^2}{\partial r}\bigg|_{r_{DNA}}-\epsilon_1\frac{\partial A_{mn}^1}{\partial r}\bigg|_{r_{DNA}}=-2\frac{\lambda_{DNA}}{r_{DNA}}$$
$$\epsilon_3\frac{\partial A_{mn}^3}{\partial r}\bigg|_{r_{protein}}-\epsilon_2\frac{\partial A_{mn}^2}{\partial r}\bigg|_{r_{protein}}=-2\frac{\lambda_{protein}}{r_{protein}} \quad \text{(A28)}$$
$$\epsilon_4\frac{\partial A_{mn}^4}{\partial r}\bigg|_{r_{surface}}-\epsilon_3\frac{\partial A_{mn}^3}{\partial r}\bigg|_{r_{surface}}=-2\frac{\lambda_{surface}}{r_{surface}}$$

For a single helix of point charges, only one of the three λs does not vanish. Thus, to evaluate V_{DNA}, the potential caused by a single DNA helix, we set $\lambda_{protein} = \lambda_{surface} = 0$. If a second DNA strand is needed, its potential may be calculated by using appropriate offsets ϕ_0 and z_0; the offsets can be factored into the original a_{mn}^i and b_{mn}^i coefficients, because both strands share the same helix parameters. To calculate the potential created by a helix of protein point charges, we set $\lambda_{DNA} = \lambda_{surface} = 0$. If a subunit contains several charges, then the net potential field is calculated by adding the contributions of all the charges, absorbing appropriate ϕ and z offsets into the coefficients. A potential caused by charges at the virion surface is calculated by setting $\lambda_{DNA} = \lambda_{protein} = 0$; each of several charges will again require its own offsets. This potential can be added to that caused by protein charges at the DNA-protein interface, because they all possess the same helix parameters. Each possible resulting system of six equations for the a_{mn}^i and the b_{mn}^i is easily inverted via a linear equations subroutine.

The simplest case has ϵ and κ constant throughout the entire space, with a single charged helix at radius R, dividing the space into region 1, interior to R, and region 2, exterior to R. This case yields

$$V_1(r)=\left(\frac{2\lambda}{\epsilon}\right)\sum_{m=-\infty}^{m=+\infty}\sum_{n=-\infty}^{n=+\infty}I_n[r\sqrt{k_{mn}^2+\kappa^2/\epsilon}]\,K_n[R\sqrt{k_{mn}^2+\kappa^2/\epsilon}]$$
$$\times\exp\{i[n(\phi-\phi_0)+k_{mn}(z-z_0)]\}$$

and (A29)

$$V_2(r)=\left(\frac{2\lambda}{\epsilon}\right)\sum_{m=-\infty}^{m=+\infty}\sum_{n=-\infty}^{n=+\infty}K_n[r\sqrt{k_{mn}^2+\kappa^2/\epsilon}]\,I_n[R\sqrt{k_{mn}^2+\kappa^2/\epsilon}]$$
$$\times\exp\{i[n(\phi-\phi_0)-k_{mn}(z-z_0)]\}.$$

In general, the product of I and K is large if their arguments are small, so the m and n values that generate the largest A_{mn} values are those for which k_{mn} is small, or $m/n \sim \Delta\phi/2\pi$. If the charge density has N_R-fold rotational symmetry, then Q is replaced by $N_R Q$, and $A_{mn}(r) \neq 0$ unless n is a multiple of N_R.

Generalizing the A_{mn} dependence

The dependence of the A_{mn} on k_{mn}^2 does not stem from the LPB equation, but from a helical distribution of point sources in a uniform medium. Suppose a point source at the origin produces a radially symmetrical potential at position $\vec{r}$ given by a Fourier integral

$$f(\vec{r})=f(r)=\int A(k^2)e^{-i\vec{k}\cdot\vec{r}}\,d^3k, \quad \text{(A30)}$$

where the wave vector $\vec{k}$ is (λ, α, ξ) in cylindrical coordinates, and A depends only on the magnitude of $\vec{k}$, because of the radial symmetry of $f(\vec{r})$. We wish to find the potential for an infinite helix of such point sources, located at positions $\vec{r}_n = (\rho, n\Delta\phi, n\Delta z)$. Equating the sum over all points to Eq. A6 gives

$$\sum_{m=-\infty}^{m=+\infty}\sum_{n=-\infty}^{n=+\infty}A_{mn}(r)\exp(in\phi)\exp(-ik_{mn}z) \quad \text{(A31)}$$
$$=\sum_{n=-\infty}^{n=+\infty}\int A(k^2)\,d^3k\exp(-i\vec{k}\cdot[\vec{r}-\vec{r}_n]).$$

The dot products can be written: $\vec{k}\cdot\vec{r}_n = \lambda\rho\cos(\alpha - n\Delta\phi) + n\,\xi\,\Delta z$ and $\vec{k}\cdot\vec{r} = \lambda r\cos(\alpha-\phi) + \xi z$. The right-hand side can be evaluated in terms of Bessel functions, expressed as (Abramowitz and Stegun, 1964)

$$J_n(x)=1/(2\pi i^n)\int_0^{2\pi}d\phi\exp(ix\cos\phi-in\phi). \quad \text{(A32)}$$

To do this, first multiply both sides of Eq. A31 by $e^{-il\phi}$ and integrate from 0 to 2π to find

$$\sum_{m=-\infty}^{m=+\infty}A_{ml}\exp(-ik_{ml}z)=(-i^l)\int d^3k\,A(k^2) \quad \text{(A33)}$$
$$\times\sum_{n=-\infty}^{n=+\infty}\exp\{i[\lambda\rho\cos(\alpha-n\Delta\phi)+n\xi\Delta z]\}\exp\{-i[\xi z+l\alpha]\}J_l(\lambda r)$$

With $d^3k = d\alpha d\xi\lambda d\lambda$, the $d\alpha$ angular integral gives a second Bessel

function:

$$\sum_{m=-\infty}^{m=+\infty} A_{ml}\exp(-ik_{ml}z) = 2\pi \int \lambda\, d\lambda J_1(\lambda r) J_1(\lambda\rho) \times\left\{\int d\xi \exp(-i\xi z) A(\lambda^2+\xi^2) \sum_{n=-\infty}^{n=+\infty} \exp(-in[l\Delta\phi - \xi\Delta z])\right\} \quad \text{(A34)}$$

The integral in the bracket can be evaluated after transforming the sum over n into a train of delta functions via Eq. A16: $\{\cdots\} = (2\pi/\Delta z)e^{-ik_{nl}z} A(\lambda^2 + k_{nl}^2)$. Substituting this into Eq. A34 and equating term-by-term yields the final result:

$$A_{nl} = \frac{(2\pi)^2}{\Delta z} \int \lambda\, d\lambda A(\lambda^2 + k_{nl}^2) J_1(\lambda r) J_1(\lambda\rho). \quad \text{(A35)}$$

Thus, the A_{nl} values result from the λ integral over the original Fourier coefficients $A(\tilde{k})$. If $f(r)$ falls off with r, then A also falls off with its argument, so in general the value of A_{nl} is maximum for a $\Delta\phi$, which causes k_{nl} to vanish.

REFERENCES

Abramowitz, M., and I. Stegun. 1964. *Handbook of Mathematical Functions.* National Bureau of Standards Applied Mathematics Series 55, Washington, D.C. 374.

Banner, D., C. Nave, and D. A. Marvin. 1981. Structure of the protein and DNA in fd filamentous bacterial virus. *Nature.* 289:814–816.

Beck, E., and B. Zink. 1981. Nucleotide sequence and genome organization of filamentous bacteriophages fl and fd. *Gene.* 16:35–58.

Berkowitz, S. A., and L. A. Day. 1976. Mass, length, composition and structure of the filamentous bacterial virus fd. *J. Mol. Biol.* 102: 531–547.

Casadevall, A., and L. A. Day. 1982. DNA packing in the filamentous viruses fd, Xf, Pf1 and Pf3. *Nucleic Acids Res.* 10:2467–2481.

Casadevall, A., and L. A. Day. 1983. Silver and mercury probing of deoxyribonucleic acid structures in the filamentous viruses fd, If1, IKe, Xf, Pf1, and Pf3. *Biochemistry.* 22:4831–4842.

Chen, F. C., G. Koopmans, R. L. Wiseman, L. A. Day, and H. L. Swinney. 1980. Dimensions of Xf virus from its rotational and translational diffusion coefficients. *Biochemistry.* 19:1373–1376.

Cross, T. A., P. Tsang, and S. J. Opella. 1983. Comparison of protein and deoxyribonucleic acid backbone structures in fd and Pf1 bacteriophages. *Biochemistry.* 22:721–726.

Day, L. A., C. J. Marzec, S. A. Reisberg, and A. Casadevall. 1988. DNA packing in filamentous bacteriophages. *Ann. Rev. Biophys. Biophys. Chem.* 17:509–539.

Day, L. A., R. L. Wiseman, and C. J. Marzec. 1979. Structure models for DNA in filamentous viruses with phosphates near the center. *Nucleic Acids Res.* 7:1393–1403.

Dunker, A. K., R. D. Klausner, D. A. Marvin, and R. L. Wiseman. 1974. Filamentous bacterial viruses. X. X-ray diffraction studies of the R4 A-protein mutant. *J. Mol. Biol.* 82:115–117.

Frank, H., and L. A. Day. 1970. Electron microscopic observations on fd bacteriophage, its alkali denaturation products and its DNA. *Virology.* 42:144–154.

Hill, D. F., N. J. Short, R. N. Perham, and G. B. Petersen. 1991. DNA sequence of the filamentous bacteriophage Pf1. *J. Mol. Biol.* 218: 349–364.

Hoffmann-Berling, H., D. A. Marvin, and H. Duerwald. 1963. Ein fadiger DNA-Phage (fd) und ein sphaerischer RNS-Phage (fr), wirtsspezifisch fuer maennliche Staemme von *E. coli.* Praeparation und chemische Eigenschaften von fd und fr. *Z. Naturforsch.* 18:876–883.

Holmes, K. C., and D. M. Blow. 1966. The use of X-ray diffraction in the study of protein and nucleic acid structure. *In* Methods of Biochemical Analysis, Vol 13. D. Glick editor. John Wiley & Sons, New York. 113–239.

Hunter, G. J., D. H. Rowitch, and R. N. Perham. 1987. Interactions between DNA and coat protein in the structure and assembly of filamentous bacteriophage fd. *Nature.* 327:252–254.

Klug, A., F. H. C. Crick, and H. W. Wyckoff. 1958. Diffraction by helical structures. *Acta Crystallogr.* 11:199–213.

Kostrikis, L. G., D. J. Liu, and L. A. Day. 1994. Ultraviolet absorption and circular dichroism of Pf1 virus: nucleotide/subunit ratio of unity, hyperchromic tyrosines and DNA bases, and high helicity in the subunits. *Biochemistry.* 33:1694–1703.

Luiten, R. G. M., D. G. Putterman, J. G. G. Schoenmakers, R. N. H. Konings, and L. A. Day. 1985. Nucleotide sequence of the genome of Pf3, an IncP-1 plasmid specific filamentous phage of Pseudomonas aeruginosa. *J. Virol.* 56:268–276.

Makowski, L., and D. L. D. Caspar. 1978. Filamentous bacteriophage Pf1 has 27 subunits in its axial repeat. *In* The Single-Stranded DNA Phages. D. T. Denhardt, D. Dressler, and D. S. Ray editors. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Makowski, L., D. L. D. Caspar, and D. A. Marvin. 1980. Filamentous bacteriophage Pf1 structure determined at 7 A resolution by refinement of models for the α-helical subunit. *J. Mol. Biol.* 140:149–181.

Makowski, L., and D. L. D. Caspar. 1981. The symmetries of filamentous phage particles. *J. Mol. Biol.* 145:611–617.

Makowski, L. 1993. Structural constraints on the display of foreign peptides on filamentous bacteriophages. *Gene.* 128:5–11.

Marvin, D. A., W. J. Pigram, R. L. Wiseman, E. J. Wachtel, and F. J. Marvin. 1974. Filamentous bacterial viruses XII. Molecular architecture of the Class I (fd, If1, IKe) virion. *J. Mol. Biol.* 88:581–600.

Marvin, D. A., R. L. Wiseman, and E. Wachtel. 1974. Filamentous bacterial viruses XI. Molecular architecture of Class II (Pf1, Xf) virion. *J. Mol. Biol.* 82:121–138.

Marvin, D. A., C. Nave, M. Bansal, R. D. Hale, and E. K. Salje. 1992. Two forms of Pf1 *Inovirus*: x-ray diffraction studies on a structural phase transition and a calculated libration normal mode of the asymmetric unit. *Phase Transitions.* 39:45–80.

Marzec, C. J., and L. A. Day. 1983. DNA and protein lattice-lattice interactions in the filamentous bacteriophages. *Biophys. J.* 42:171–180.

Marzec, C. J., and L. A. Day. 1988. A theory of the symmetries of filamentous bacteriophages. *Biophys. J.* 53:425–440.

Marzec, C. J., and L. A. Day. 1993. Pattern formation in icosahedral virus capsids: the papovaviruses and Nudaurelia capensis β virus. *Biophys. J.* 65:2559–2577.

Model, P., and M. Russel. 1988. Filamentous bacteriophage. *In* The Bacteriophages. R. Calendar editor. Plenum, New York.

Nave, C., A. G. Fowler, J. E. Ladner, D. A. Marvin, S. W. Provencher, A. Tsugita, J. Armstrong, and R. N. Perham. 1981. Pf1 filamentous bacterial virus. X-ray fibre diffraction analysis of two heavy-atom derivatives. *J. Mol. Biol.* 149:675–707.

Newman, J., H. L. Swinney, and L. A. Day. 1977. Hydrodynamic properties and structure of fd virus. *J. Mol. Biol.* 116:593–603.

Newman, J., L. A. Day, G. Dalack, and D. Eden. 1982. Hydrodynamic determination of molecular weight, dimensions, and structural parameters of Pf3 virus. *Biochemistry.* 21:3352–3358.

Peterson, C., W. T. Winter, G. Dalack, and L. A. Day. 1982. Structure of the filamentous bacteriophage Pf3 by X-ray fiber diffraction. *J. Mol. Biol.* 162:877–881.

Reisberg, S. A. 1989. Radial mass density profiles, mass per length, and nucleotide/subunit ratios for filamentous bacteriophages from electron microscopy, Ph.D. thesis. New York University, New York.

Soumpasis, D. 1978. Debye-Huckel theory of model polyelectrolytes. *J. Chem. Phys.* 69:3190–3196.

Wachtel, E. J., R. L. Wiseman, W. J. Pigram, and D. A. Marvin. 1974. Filamentous bacterial viruses XIII. Molecular structure of the virion in projection. *J. Mol. Biol.* 88:601–618.

Wiseman, R. L., S. A. Berkowitz, and L. A. Day. 1976. Different arrangements of protein subunits and single-stranded circular DNA in the filamentous bacterial viruses fd and Pf1. *J. Mol. Biol.* 102:549–561.

Wiseman, R. L., and L. A. Day. 1977. Different packaging of DNA in the filamentous viruses Pf1 and Xf. *J. Mol. Biol.* 116:607–611.